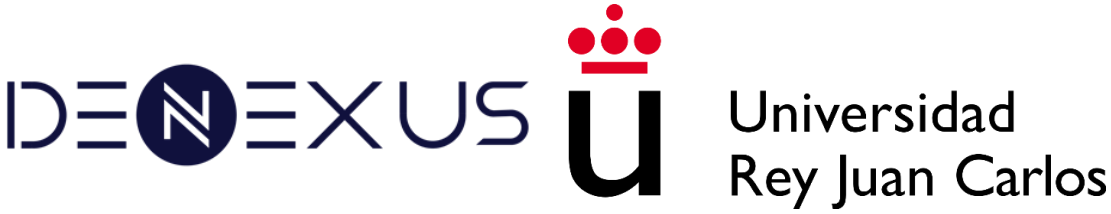
Convocatoria Colaboración Público-Privada CPP2021-009025

DICYME:

Dynamic Industrial CYberrisk Modelling based on Evidence

Iniciativa Conjunta de:



ENTREGABLE 3.2:
Primer prototipo de interfaz gráfica y herramienta de soporte a la decisión

Coordinadores:

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

Contenido

1 INTRODUCCIÓN Y OBJETIVOS

Este documento resume los pasos realizados en el desarrollo de la Actividad 3 (Sistema de Visualización y toma de decisiones), con el resultado del segundo prototipo de interfaz gráfica y herramienta de soporte, desarrollado sobre la base del primer prototipo (entregable E3.1).

Este sistema de visualización se vale de los datos extraídos gracias a los módulos de extracción automática de datos desarrollados en la Actividad 1 (Sistema de Recogida de evidencias) y a los resultados de los módulos de medida/estimación de probabilidad e impacto de la Actividad 2 (Módulos de medida/estimación de probabilidad e impacto).

En las siguientes secciones se explica:

- Cómo se ha construido la nueva navegación de la aplicación para guiar al usuario en el proceso desde la adquisición de los datos hasta la cuantificación del ciberriesgo.
- Las mejoras que se han introducido en las funcionalidades presentes en el primer prototipo (entregable E3.1).
- Las nuevas funcionalidades y visualizaciones se han añadido con los resultados del proyecto para facilitar la toma de decisiones.
- Los datos y código que se encuentran disponibles en este entregable (deriskGroup / DICyME Project · GitLab) y cómo pueden emplearse.

2 CAMBIOS Y MEJORAS DEL SISTEMA DE VISUALIZACIÓN

2.1 Nueva navegación

La aplicación descrita a lo largo de este entregable corresponde con la versión v0.4.0, cuyo desarrollo concluyó en octubre de 2024.

Aunque se han mantenido los grupos definidos en el menú principal con respecto al primer prototipo, se han renombrado y reordenado para facilitar al usuario la comprensión del proceso seguido con los datos, así como las posibilidades que ofrece la plataforma.

De esta manera, el usuario primeramente accede a una pestaña principal, que actualmente tiene una fotografía estática con el nombre y referencia del proyecto.

Seguidamente, pasa al módulo de datos, en el que se mantienen las pestañas de ciber incidentes y de perfil de víctimas (entregable E1.2). Se mantienen pestañas en las que se planean añadir más fuentes de datos, pero que no han sido desarrolladas aún.

Una vez ha analizado los datos, el usuario puede moverse al módulo de Indicadores, donde se muestran resultados de los modelos y algoritmos desarrollados en el proyecto.

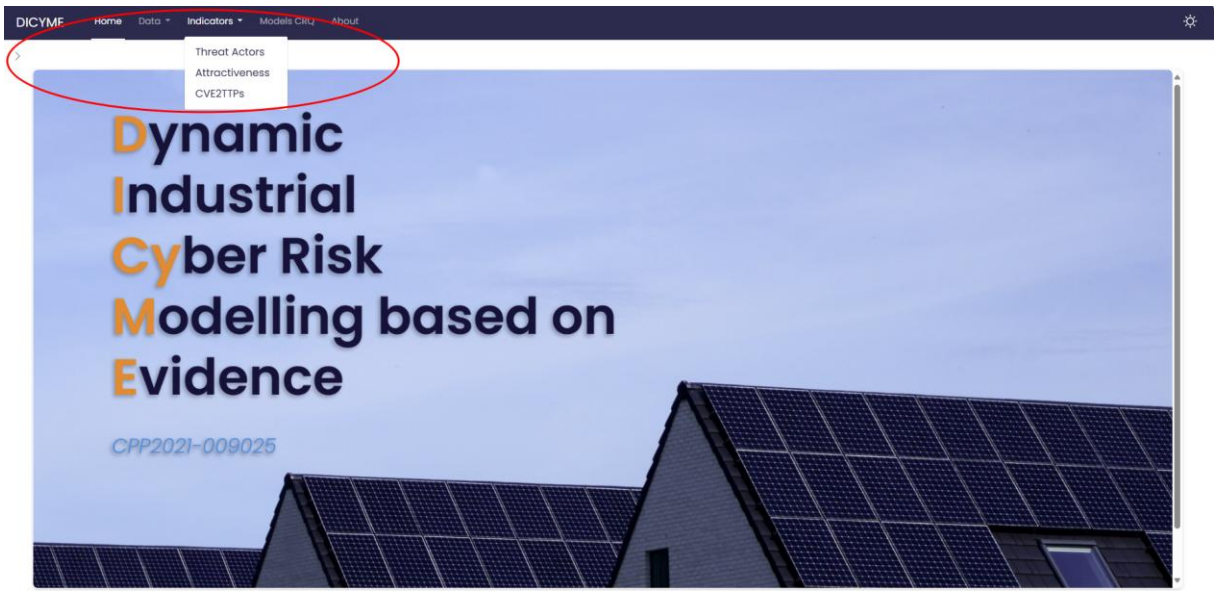


Ilustración 1. Pestaña principal y nuevo menú de navegación

Y una vez expuesto todo lo anterior, por último, el usuario puede navegar al módulo de modelos de cuantificación del ciberriesgo (Modelo CRQ), en el que se combinan los datos y algoritmos previos para cuantificar el ciberriesgo de una entidad. Este apartado forma parte de la Actividad 4 (Integración, validación y evaluación) del proyecto.

Adicionalmente, permanece la pestaña de información sobre el proyecto y las fuentes de datos. El nuevo menú puede visualizarse en la Ilustración 1. Pestaña principal y nuevo menú de navegación.

2.2 Módulo de datos

Dentro de este módulo, se ha mejorado la visualización de las dos fuentes de datos ya añadidas en el primer prototipo: ciber incidentes y perfil de víctimas (conjunto de ciber incidentes enriquecido).

La primera mejora corresponde con los gráficos de series temporales, que anteriormente eran gráficos de barras. Para los ciber incidentes, se muestran series temporales independientes, puesto que puede haber incidentes presentes en múltiples bases de datos simultáneamente (ver Ilustración 2. Serie temporal de ciber incidentes.). Para el conjunto de datos del perfil de víctimas, la serie temporal es agregada, puesto que no existe dicha duplicación (ver Ilustración 3. Serie temporal de ciber incidentes enriquecidos.).

Para los ciber incidentes, se ha añadido un nuevo gráfico que muestra, para diferentes variables y para cada base de datos en la que existen dichos datos, el recuento de cada valor existente agrupado por la unidad de tiempo seleccionada en el menú izquierdo (ver Ilustración 4. Gráfico de barras para diferentes variables de los ciber incidentes.).

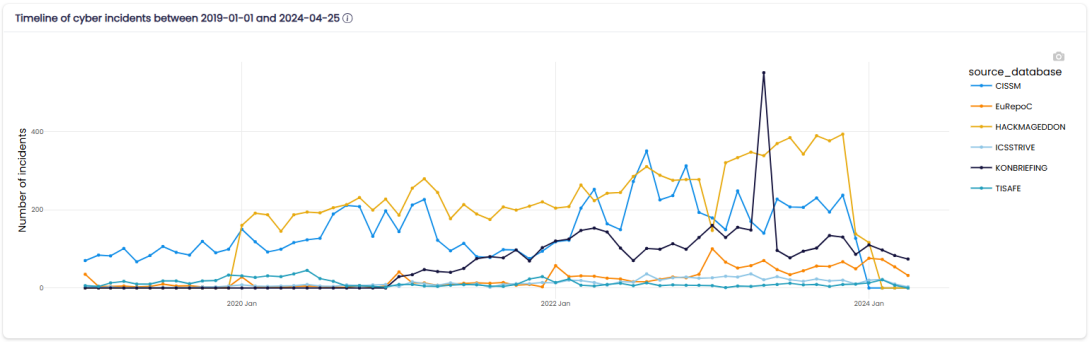


Ilustración 2. Serie temporal de ciber incidentes.

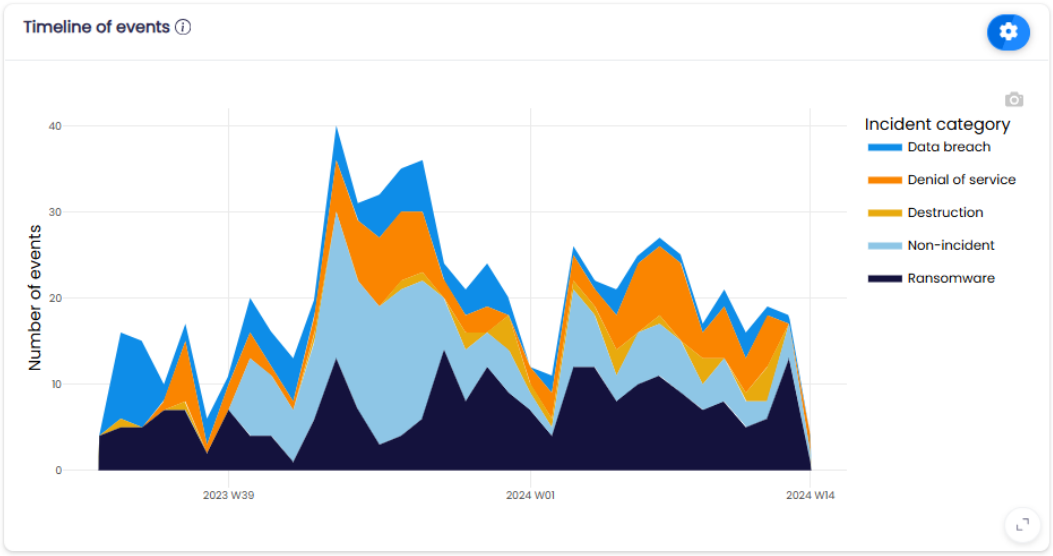


Ilustración 3. Serie temporal de ciber incidentes enriquecidos.

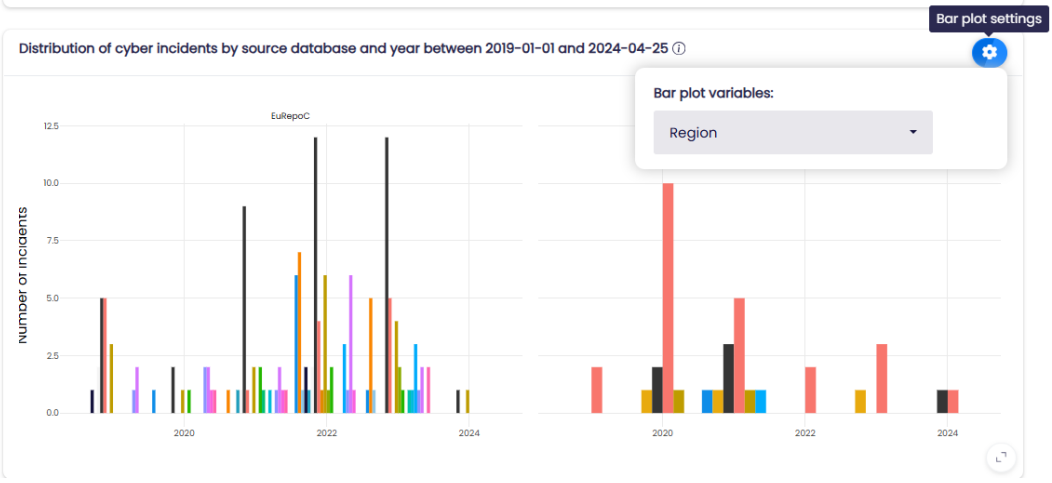


Ilustración 4. Gráfico de barras para diferentes variables de los ciber incidentes.

También se ha mejorado la visualización de los datos en formato tabla, eliminando los valores nulos, formateando las cantidades económicas, los valores lógicos/booleanos, cifras decimales, etc. (ver Ilustración 5. Tabla con datos de ciber incidentes enriquecidos.).

	Date	Entity	Incident category	Country	Category	Revenue	Earnings	Employees	Profitable
	All	All	All	All	All	All	All	All	All
1	2023-08-20	.au Domain Administration (auDA)	Ransomware	Australia	Organizations			70	No
3	2023-12-13	3T Software Labs GmbH	Non-incident	Germany	Software	4,000,000€		74	Yes
4	2024-02-06	3g office by 3g Smart Group	Non-incident	Spain	Business Services	23,000,000€		182	Yes
5	2023-10-11	3rd Millennium Classrooms	Data breach	United States	Education	3,000,000€		28	Yes
6	2024-02-15	AB Texel	Ransomware	United Kingdom	Agriculture & Fishing	3,000,000€		1	Yes
7	2023-11-12	ACFS Port Logistics	Non-incident	Australia	Trade, Supply Chain & Commerce	3,800,000€		355	Yes
8	2023-12-25	ALBtelecom (Eagle Mobile)	Destruction	Albania	Telecommunications	234,000,000€		748	Yes
9	2023-10-09	AMPLIA Communications	Non-incident	Trinidad and Tobago	Telecommunications			174	Yes

Ilustración 5. Tabla con datos de ciber incidentes enriquecidos.

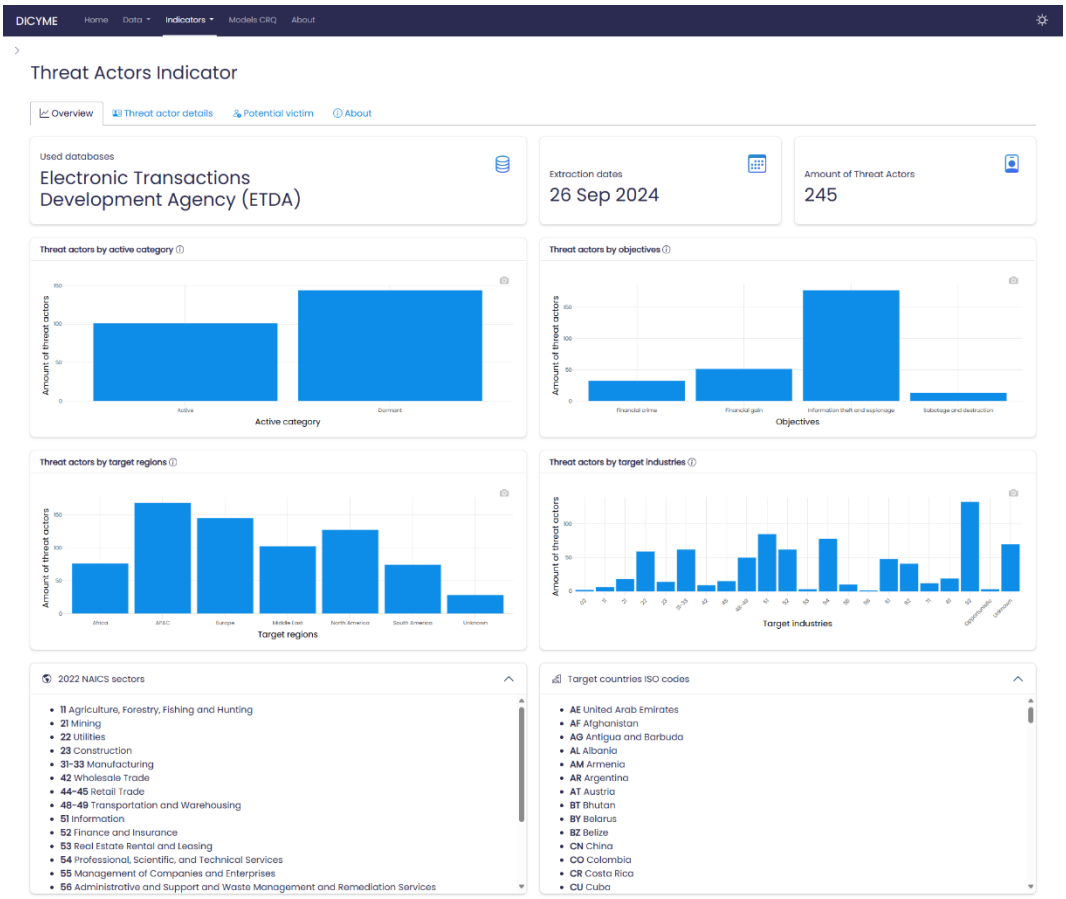


Ilustración 6. Página principal del indicador de actores de amenazas.

2.3 Módulo de indicadores

Este segundo módulo comprende la mayor parte del trabajo desarrollado desde el primer prototipo. En primer lugar, se ha desarrollado la pestaña de visualización del indicador de actores de amenaza, empleando para su cálculo datos de la *Electronic Transactions Development Agency* (ETDA). Al acceder, inicialmente se muestran estadísticas agregadas de todos los actores, como la cantidad de estos por categoría de actividad, objetivos, países a los que atacan, etc. También se incluye un glosario con la nomenclatura y códigos de las industrias según la *North American Industry*

Ilustración 7. Página de detalles de actores de amenazas.

Classification System (NAICS) y los códigos ISO de los países presentes en los datos (ver Ilustración 6. Página principal del indicador de actores de amenazas.).

El usuario puede navegar seguidamente a una pestaña que muestra información detallada de un actor de amenazas concreto que puede escoger en el menú de selección. En este lugar puede comprobar el valor final del indicador para diferentes combinaciones de países e industrias, así como el valor medio del indicador, y las puntuaciones parciales generales asignadas a la actividad y capacidad del actor con un código de colores por si es elevado —naranja si es mayor que 0.5—, o no —turquesa— (ver Ilustración 7. Página de detalles de actores de amenazas.).

Por último, una tercera página muestra, para una víctima ficticia definida por el país e industria (seleccionables en el menú), la cantidad de actores involucrados, así como la puntuación de cada uno de ellos. Un gráfico de barras circular combina el indicador de actores con las puntuaciones de actividad y capacidad para los 20 actores con valor de indicador más alto (ver Ilustración 8. Páginas de actores de amenazas para una víctima potencial).

La siguiente pestaña corresponde al cálculo de atractivo de una entidad, ofreciendo una interfaz que permite la comparación entre hasta 3 entidades. Tras seleccionar en el desplegable las entidades deseadas, se muestra su valor de atractivo final, así como la representación de sus variables numéricas de manera normalizada en un gráfico de radar, de manera que se puedan comparar las diferencias entre las mismas. También se

muestran todas las variables de las entidades en formato tabla y con los valores reales (ver Ilustración 9. Pestaña de atractivo).

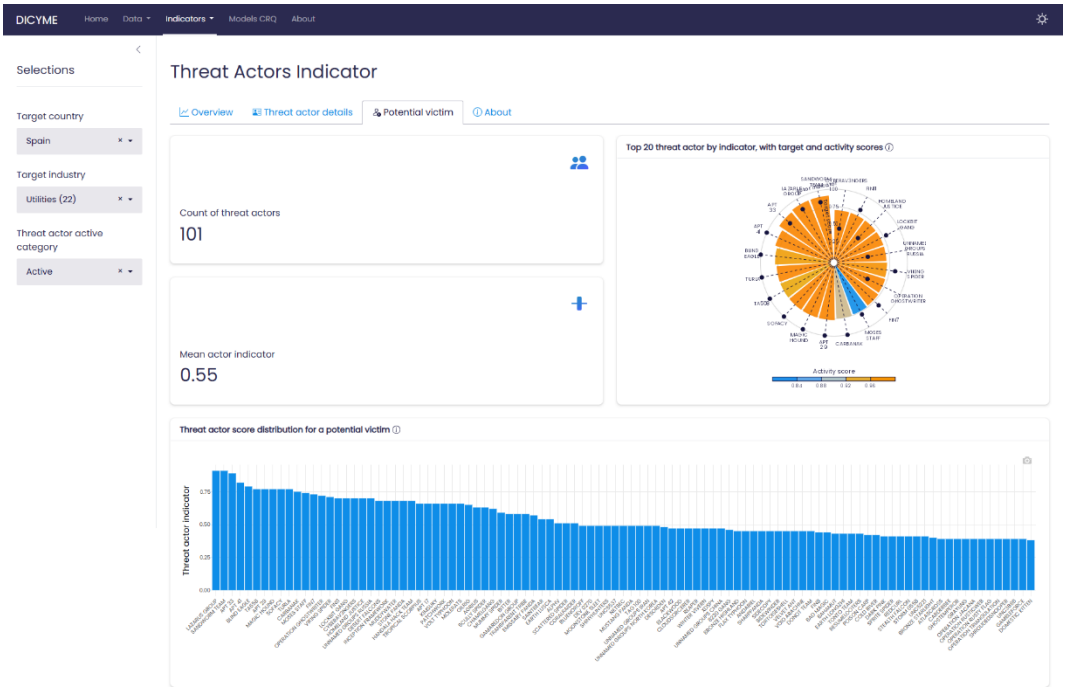


Ilustración 8. Páginas de actores de amenazas para una víctima potencial

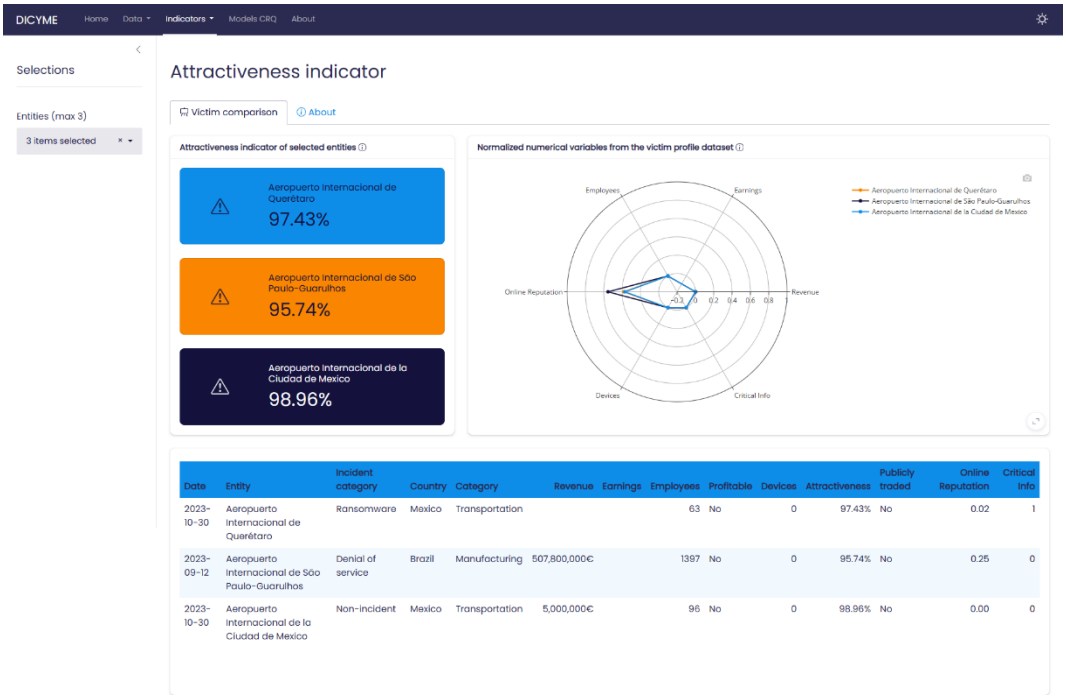


Ilustración 9. Pestaña de atractivo

La tercera y última pestaña desarrollada está relacionada con el modelo CVE2TTPs, que relaciona vulnerabilidades de la *Common Vulnerabilities and Exposures* (CVE) con las tácticas, técnicas y procedimientos (TTPs) de las matrices empresarial (*Enterprise*) y de sistemas de control industrial (*Industrial Control Systems*, *ICS*) de MITRE ATT&CK. De

momento, toda la información mostrada en esta pestaña es estática y no hay selectores que posibiliten el dinamismo de otras pestañas.

Ilustración 10. Pestaña del modelo CVE2TTPs.

Para cada una de las dos matrices, se ha desarrollado un mapa de calor que muestra, mediante una escala de colores, la cantidad de vulnerabilidades predichas para cada par táctica-tipo de vulnerabilidad. También se ha añadido un gráfico de barras mostrando la cantidad de vulnerabilidades y la cantidad de técnicas de cada matriz predichas por año. El último gráfico, de tipo pirámide, muestra el valor medio del *Common Vulnerability Scoring System* (CVSS), un sistema de puntuación que estima el impacto de una vulnerabilidad, tanto en su versión v2 como v3. A cada lado de la pirámide se muestra

cada una de las matrices: *Enterprise* e *ICS* (ver Ilustración 10. Pestaña del modelo CVE2TTPs.).

3 DOCUMENTACIÓN Y PRUEBAS

Como novedad respecto al primer prototipo, se ha añadido documentación de todas las funciones definidas en el código siguiendo el formato *roxygen*, de manera que la aplicación sea más fácil de mantener y entender por otros desarrolladores. **Roxygen** es una herramienta que permite escribir la documentación directamente en el código fuente, lo que facilita mantener la documentación actualizada y coherente con el desarrollo.

Además, con toda la documentación, así como el archivo NEWS.md, que contiene los avances y novedades de cada versión, se genera un sitio web interactivo y consistente con multitud de paquetes haciendo uso de pkgdown. Dicho sitio web incluye los cambios en cada versión, la documentación de las funciones, los ejemplos ejecutados junto a su salida, entre otros, facilitando su consulta a desarrolladores y, por tanto, asegurando la comprensión y mantenimiento del código. Este sitio web se encuentra desplegado junto a la propia aplicación en la siguiente ruta:

https://gondor.etsii.urjc.es:3866/www/docs/index.html

Por otro lado, también se han añadido pruebas con *testthat* que comprueban los tipos de los objetos devueltos por todas las funciones del paquete. *testthat* es una de las herramientas más populares para realizar pruebas unitarias en R, y su uso garantiza que las funciones del paquete se comporten como se espera, mejorando la fiabilidad y la calidad del código.

En resumen, estas mejoras no solo optimizan la presentación y accesibilidad de la documentación, sino que también aseguran que el código sea robusto y fácil de mantener, facilitando la colaboración y el desarrollo continuo del paquete.

4 DATOS Y CÓDIGO

En cuanto al desarrollo y despliegue no se ha modificado el flujo de trabajo definido desde el comienzo y detallado en el entregable E3.1. La aplicación es accesible en la URL anteriormente indicada y se actualiza en intervalos cortos de desarrollo, generalmente de dos semanas, introduciendo las nuevas funcionalidades que se van desarrollando, así como generando el paquete de R con el código fuente para cada versión.

En el enlace deriskGroup / DICyME Project · GitLab, dentro del directorio *E.3.2 First prototype of a graphical interface and decision support tool*, se puede encontrar el directorio `dicymeviz-0.4.0/` que contiene el código de la aplicación en formato paquete de R. El archivo README.md detalla la estructura y modo de ejecución de la aplicación.

De igual manera que en el entregable E3.1, caben destacar algunos elementos del código:

- `renv.lock` contiene las versiones de todos los paquetes de R empleados, así como sus dependencias y versiones de las mismas.
- `R/` contiene el código R con las diferentes funciones que componen la aplicación.
- `.github/workflows/` contiene el código con las automatizaciones del despliegue y análisis estático.
- DESCRIPTION contiene la información del paquete, versión, autores, etc.

Aunque las versiones de todas las dependencias aparecen especificadas en el archivo renv.lock y se pueden restaurar de manera sencilla en cualquier entorno de desarrollo, cabe destacar las siguientes:

- `R 4.4.1`
- `Shiny 1.9.1`
- `bslib 0.8.0.9000`

5 CONCLUSIONES Y SIGUIENTES PASOS

En este punto del proyecto se ha desarrollado un segundo prototipo de una aplicación web mediante la tecnología Shiny en R. Tras el primer prototipo, en este se ha mejorado la visualización de las fuentes de datos previamente añadidas, modificando los gráficos y añadiendo otros nuevos. Además, se han añadido visualizaciones para los resultados de los modelos desarrollados en el proyecto, mostrando la información útil que son capaces de generar, y por tanto poniendo de manifiesto el valor que aportan en la toma de decisiones.

En el siguiente prototipo se pretende añadir las últimas fuentes de datos empleadas en el proyecto, así como nuevos selectores que añadan dinamismo y control sobre la información mostrada en las diferentes pestañas. También se prevé añadir un informe personalizado con la información extraída de los modelos, de manera que se facilite la toma de decisiones y la difusión de los datos e información que soportan la misma, por ejemplo, para compartirlo entre responsables del proceso, entre técnicos y directivos, o entre diferentes técnicos encargados de la remediación.

Además, se trabajará en la posibilidad de permitir la ejecución de los modelos de estimación de probabilidad e impacto, ya que actualmente tan sólo se muestran

resultados de estos para unas entradas definidas. Esto permitirá la generación de múltiples escenarios diferentes adaptados a cada caso de uso y organización concretos, pudiendo así facilitar más aún la toma de decisiones con simulaciones.

Por último, se planea trabajar en profundidad la dimensión de explicabilidad de los resultados, así como la recomendación de modificaciones factibles de las entradas de los modelos que las entidades o sujetos podrían realizar para cambiar las predicciones y mejorar por tanto su postura de ciberseguridad —reducir su ciber riesgo—.