



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



**TFG del Grado en Ingeniería
Informática**

**Sistema de información de
apoyo a GreenMetrics
Documentación Técnica**



Presentado por Lorena Bueno Porras
en Universidad de Burgos — 6 de junio
de 2025

Tutor: José Manuel Galán Ordax y Virginia
Ahedo García

Índice general

Índice general	i
Índice de figuras	iii
Índice de tablas	v
Apéndice A Plan de Proyecto Software	1
A.1. Introducción	1
A.2. Planificación temporal	1
A.3. Estudio de viabilidad	19
Apéndice B Especificación de Requisitos	25
B.1. Introducción	25
B.2. Objetivos generales	25
B.3. Catálogo de requisitos	25
B.4. Especificación de requisitos	28
Apéndice C Especificación de diseño	43
C.1. Introducción	43
C.2. Diseño de datos	43
C.3. Diseño arquitectónico	44
C.4. Diseño de interfaces	46
C.5. Diseño procedimental	46
Apéndice D Documentación técnica de programación	57
D.1. Introducción	57
D.2. Estructura de directorios	57

D.3. Manual del programador	58
D.4. Compilación, instalación y ejecución del proyecto	60
D.5. Pruebas del sistema	61
Apéndice E Documentación de usuario	63
E.1. Introducción	63
E.2. Requisitos de usuarios	63
E.3. Instalación	64
E.4. Manual del usuario	64
Apéndice F Anexo de sostenibilización curricular	67
F.1. Introducción	67
F.2. Impacto de la sostenibilidad en el proyecto	67
F.3. Competencias adquiridas y aplicadas	68
F.4. Conclusión	69
Bibliografía	71

Índice de figuras

A.1. Sprint Inicio	3
A.2. Sprint de Implementación de Descarga y Filtrado de Guías Docentes	5
A.3. Sprint de Generación de Informes Específicos	6
A.4. Sprint de Mejora de Interfaces y Automatización de Informes	8
A.5. Sprint de Ajustes en Informes: Correcciones y Nuevas Funcionalidades	9
A.6. Implementación de Filtrado Avanzado y Nuevos Informes de Sostenibilidad	11
A.7. Sprint de Mejora de Interfaz, Tabla de Búsqueda e Implementación de Nuevos Informes	13
A.8. Sprint de Informes Académicos, Gestión de Roles y Documentación Técnica	15
A.9. Sprint de Documentación y Organización del Proyecto	17
A.10. Distribución de Issues por Categoría	19
A.11. Régimen general de la Seguridad Social I.	20
A.12. Régimen general de la Seguridad Social II.	20
C.1. Patrón MVC [5]	45
C.2. Patrón fachada [6]	46
C.3. Diagrama de secuencia de Registro y Login	47
C.4. Diagrama de secuencia evidencia 1_19	48
C.5. Diagrama de secuencia de la evidencia 6_1 para consultar datos	49
C.6. Diagrama de secuencia de la evidencia 6_1 para descargar datos	50
C.7. Diagrama de secuencia de la evidencia 6_1 para generar el informe	51
C.8. Diagrama de secuencia de las evidencias 6_2 y 6_3	52
C.9. Diagrama de secuencia de la evidencia 6_4	53
C.10. Diagrama de secuencia de la evidencia 6_7	54
C.11. Diagrama de secuencia de la evidencia 6_8	55

E.1. Esquema carpetas	65
---------------------------------	----

Índice de tablas

A.1. Relación entre puntos de historia y estimación temporal	2
A.2. Horas dedicadas al proyecto	18
A.3. Costes- totales	22
A.4. Licencias de herramientas y librerías	23
A.5. Comparación de licencias	23
B.1. CU-1 Registro de usuario.	29
B.2. CU-2 Registro de login.	30
B.3. CU-3 Procesamiento de guías docentes.	31
B.4. CU-4 Automatización con modelos de IA.	32
B.5. CU-5 Gestión de base de datos	33
B.6. CU-6 Editar/eliminar filas	34
B.7. CU-7 Mostrar base de datos	35
B.8. CU-8 Configuración	36
B.9. CU-9 Configuración perfil	36
B.10.CU-10 Ayuda de manual	37
B.11.CU-11 Informe 1_19	38
B.12.CU-12 Informe 6_1	39
B.13.CU-13 Informe 6_2	39
B.14.CU-14 Informe 6_3	40
B.15.CU-15 Informe 6_4	41
B.16.CU-16 Informe 6_7	42
B.17.CU-17 Informe 6_8	42

Apéndice A

Plan de Proyecto Software

A.1. Introducción

Este apartado del anexo está dedicado al análisis de aspectos fundamentales para la ejecución del proyecto. En primer lugar, se presenta la planificación temporal, donde se establecen las diferentes fases y tareas que componen el desarrollo del trabajo. A continuación, se incluye un apartado de viabilidad económica, en el cual se estima el coste derivado de los recursos necesarios para la implementación del mismo. Finalmente, se estudia la viabilidad legal, considerando la normativa vigente que podría tener impacto en el proyecto.

A.2. Planificación temporal

Al comienzo del proyecto se propuso utilizar una metodología ágil para la gestión del proyecto. De entre todas las opciones se escogió emplear Scrum, ya que permite entre otras cosas, entregar incrementos de producto funcionales de manera regular. También permite ajustar las prioridades en cada sprint para responder a nuevas necesidades que pudieran darse. Aunque no ha podido seguirse al 100 % esta metodología debido a que no se cumplían todos los requisitos necesarios, como ser un equipo de 4 a 8 personas, hacer reuniones diarias. . . , se ha aplicado de manera general.

- Se siguió una estrategia de desarrollo incremental a través de sprints (iteraciones).
- Cada sprint tuvo una duración de dos semanas.

- Al final de cada sprint se entregaba una parte funcional del producto.
- Al finalizar cada sprint se realizaban reuniones de revisión de dicha iteración, además de planificar el nuevo sprint.
- Durante la planificación del sprint se realizaba una pila de tareas que debían completarse durante esa iteración.
- Para monitorizar el avance del proyecto se utilizaron gráficos burndown.

La tabla A.1 muestra la correspondencia estimada entre los puntos de historia y el tiempo necesario para completar cada tarea.

Puntos de historia	Estimación temporal
1	2 horas
2	10 horas
3	1 día
5	3 días
8	5 días
13	7 días
21	10 días
40	15 días

Tabla A.1: Relación entre puntos de historia y estimación temporal

Sprint(17/01/2025 – 31/01/2025)

– Inicio –

Este sprint marcó el inicio del proyecto. Los objetivos principales fueron desarrollar una interfaz funcional para las diferentes páginas (login, registro, página principal, perfil y ajustes), permitir el inicio de sesión con Google, y crear el código encargado de buscar las guías docentes de todos los grados y másteres de la UBU. Se instaló de forma local un modelo de lenguaje (LLM) que recorre los archivos PDF de las guías docentes descargadas para identificar: nombre de la asignatura, nombre del grado/máster, código y competencias sostenibles.

Aunque se estimó que la duración de este sprint sería de 15 días, debido a que quedaba una tarea por completar, el sprint se extendió a 19 días en total.

La Figura A.1 ilustra las tareas abordadas durante el sprint inicial del proyecto.

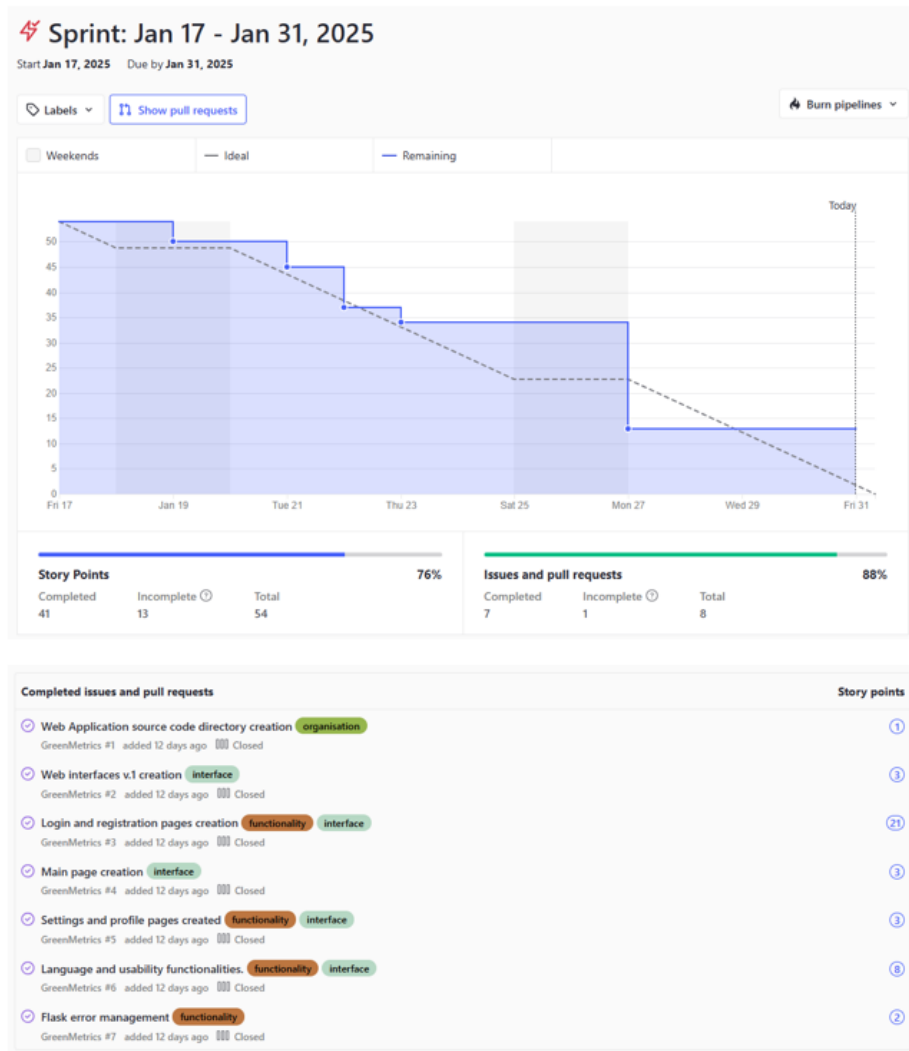


Figura A.1: Sprint Inicio

Sprint(31/01/2025 – 14/02/2025)

– Sprint de Implementación de Descarga y Filtrado de Guías Docentes –

Los objetivos principales consistieron en descargar las guías docentes de un año determinado y permitir la selección entre grado, máster o ambos.

Durante el proceso de descarga, se fue completando automáticamente la base de datos, la cual puede ser consultada y filtrada según diferentes criterios: curso académico, código de la asignatura, tipo de estudio (grado, máster o ambos), nombre del archivo y modalidad. Además, en este sprint se añadió la funcionalidad que permite al usuario seleccionar el modelo de lenguaje (LM) que prefiera, el cual recorre los archivos PDF de las guías docentes descargadas para identificar aquellas que contienen competencias. Cabe destacar que la implementación del LM ya se había completado en el sprint anterior. Finalmente, se generaron archivos PDF que cumplen con un formato especificado en un archivo DOCX. Esta parte se desarrolló de manera general, con el objetivo de poder personalizarla en el futuro, según las necesidades específicas del informe que se desee generar.

Se estimó una duración de 15 días.

En la Figura [A.2](#) se muestra la planificación del sprint centrado en la descarga y filtrado de guías docentes.

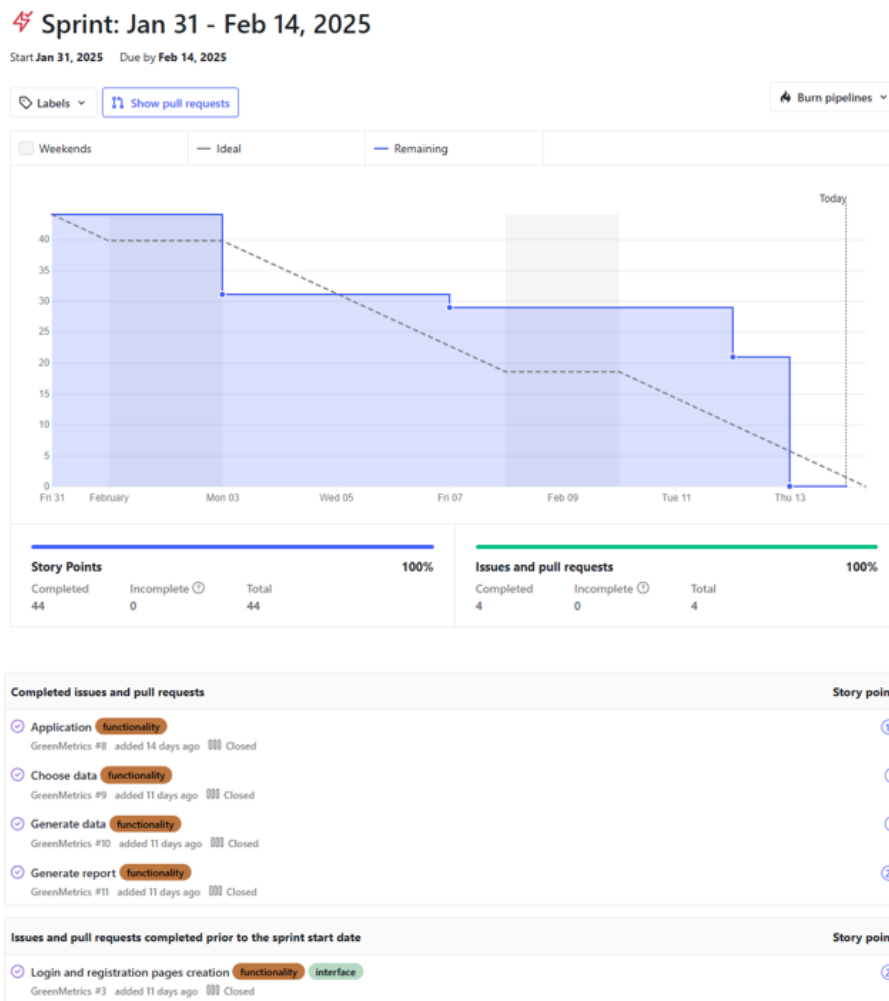


Figura A.2: Sprint de Implementación de Descarga y Filtrado de Guías Docentes

Sprint(14/02/2025 – 28/02/2025)

– Sprint de Generación de Informes Específicos –

Los objetivos principales de este sprint fueron, a partir del informe general implementado en el sprint anterior, generar dos informes específicos. El primero, "**Number of Courses on Environment and Sustainability (6_1)**", requería obtener los resultados de la API (implementada en el sprint anterior) y generar un PDF con el formato adecuado. El segundo, "**Annual Operation Maintenance Percentage (1_19)**", implicó implementar una función capaz de realizar búsquedas en la página de licitaciones de la

UBU basándose en los parámetros introducidos por el usuario. Los enlaces obtenidos fueron procesados por una API que extrae únicamente los datos necesarios para el informe, tras lo cual se generó el PDF correspondiente.

Se estimó una duración de 15 días.

La Figura A.3 representa el desarrollo del sprint destinado a la generación de informes específicos a partir de la API implementada.

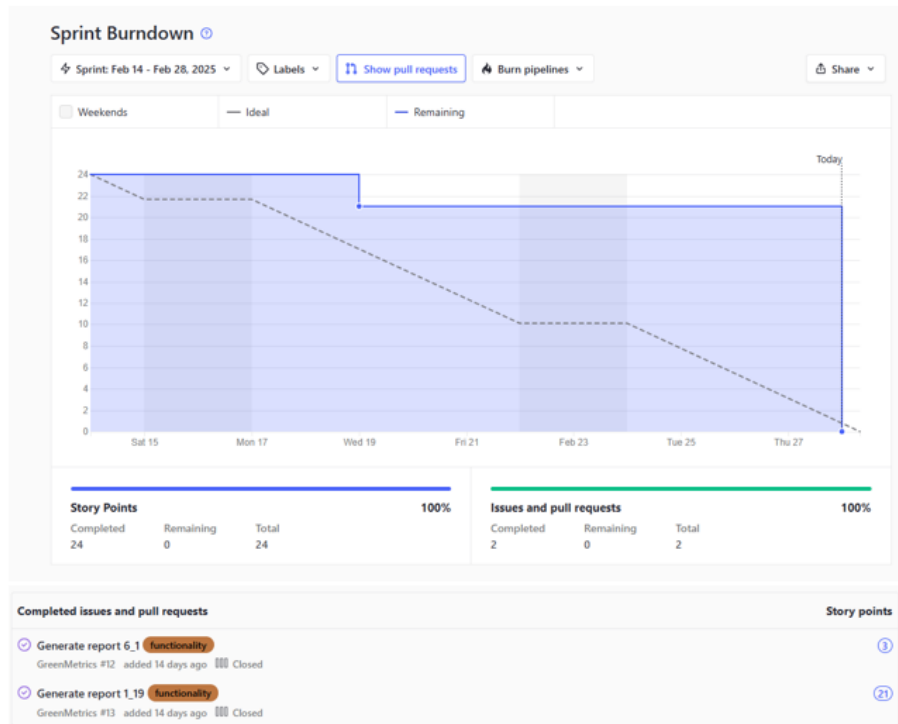


Figura A.3: Sprint de Generación de Informes Específicos

Sprint(28/02/2025 – 14/03/2025)

– Sprint de Mejora de Interfaces y Automatización de Informes –

Los objetivos principales de este sprint fueron mejorar la usabilidad de las interfaces, automatizar la búsqueda de licitaciones de mantenimiento, actualizar la clasificación de competencias en los informes y añadir un nuevo informe sobre publicaciones en sostenibilidad. En primer lugar, se reorganizaron las interfaces para que fueran más intuitivas y user-friendly. Además, se modificó el informe “**Annual Operation Maintenance Percentage (1_19)**”, de modo que, en lugar de realizar búsquedas basadas

en los parámetros introducidos por el usuario, se automatizara la búsqueda exclusivamente para licitaciones relacionadas con mantenimiento. También se actualizó el informe "**Number of Courses on Environment and Sustainability (6_1)**", permitiendo clasificar las competencias en función de la categoría de desarrollo sostenible a la que pertenecen. Asimismo, se añadió el informe "**Number of Scholarly Publications on Sustainability (6_7)**", que busca en Google Scholar el número de publicaciones sobre sostenibilidad en un rango de años especificado por el usuario. Para ello, se implementó un sistema que captura automáticamente una imagen de la búsqueda y genera el informe insertando dicha captura, además de incluir el enlace de la búsqueda y una descripción. El único inconveniente encontrado fue un problema con el enlace dentro del PDF: al convertir el archivo, el enlace insertado no funcionaba correctamente. Se prevé solucionar este error en futuros sprints.

Se estimó una duración de 15 días.

La Figura A.4 muestra la automatización de informes y mejoras en la interfaz llevadas a cabo en ese sprint.

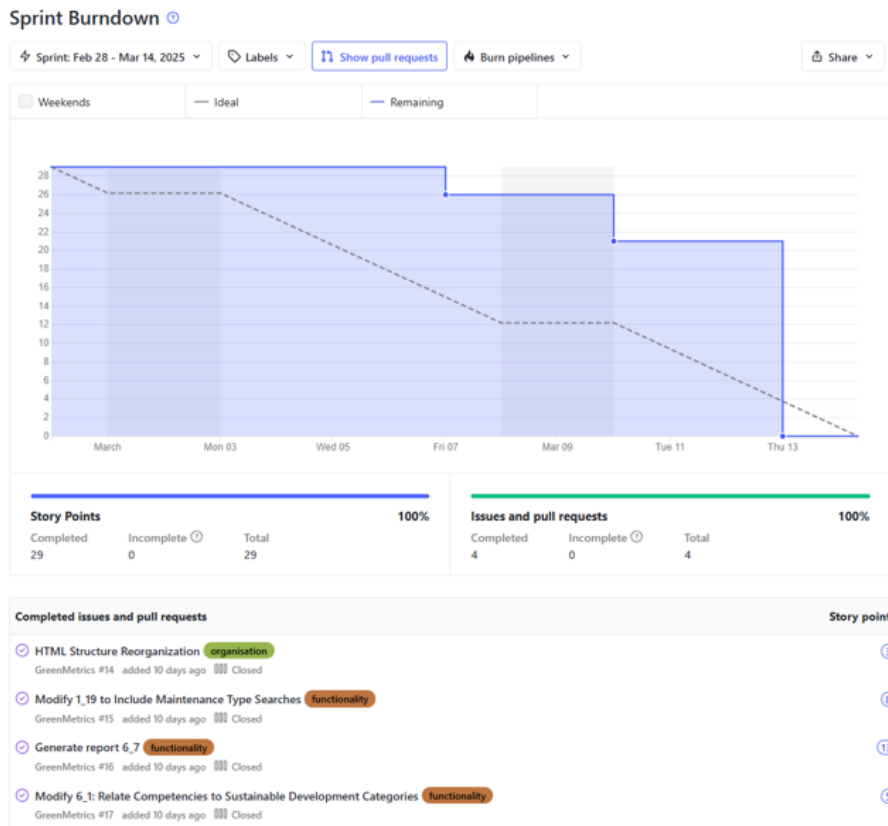


Figura A.4: Sprint de Mejora de Interfaces y Automatización de Informes

Sprint(14/03/2025 – 28/03/2025)

– Sprint de Ajustes en Informes: Correcciones y Nuevas Funcionalidades –

Los objetivos principales de este sprint fueron; corregir los errores en los enlaces de los informes; cambiar el nombre de los informes en las interfaces; modificar el informe “**Annual Operation Maintenance Percentage (1_19)**” para que también realice búsquedas por tipos de mantenimiento y generar un nuevo informe: “**Number of events related to environment and sustainability (6_8)**”. En primer lugar, se corrigieron los enlaces y se actualizaron los nombres de todos los informes. También se actualizó el código del informe “**Annual Operation Maintenance Percentage (1_19)**”, de manera que no solo busque por la palabra clave “mantenimiento”, sino que también tenga en cuenta los distintos tipos de mantenimiento. Por último, se añadió el informe “**Number of events related to environment and**

sustainability (6_8)", el cual permite al usuario seleccionar un rango de años, buscar los archivos correspondientes a dichos años y determinar cuántas actividades sostenibles se han llevado a cabo, para posteriormente generar un informe.

Aunque se estimó que la duración de este sprint sería de 15 días, debido a que quedaban tareas por completar, el sprint se extendió a 19 días en total.

En la Figura A.5 se reflejan las tareas de ajuste y creación de nuevos informes con funcionalidades ampliadas.

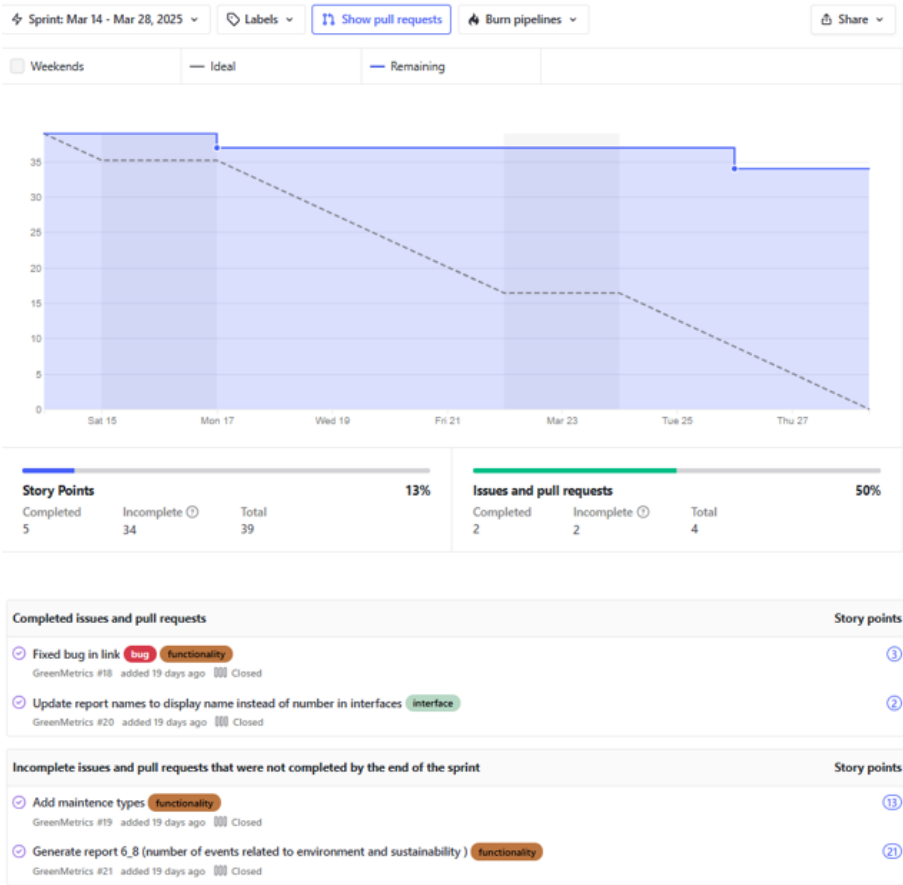


Figura A.5: Sprint de Ajustes en Informes: Correcciones y Nuevas Funcionalidades

Sprint(28/03/2025 – 11/04/2025)

–Implementación de Filtrado Avanzado y Nuevos Informes de Sostenibilidad–

Los objetivos principales de este sprint fueron añadir al informe “**Annual Operation Maintenance Percentage (1_19)**” la capacidad de filtrar por los diferentes tipos de mantenimiento, así como generar dos nuevos informes: “**Number of events related to environment and sustainability (6_8)**” y “**Total research funds dedicated to sustainability research (in US dollars) (6_4)**”, junto con sus correspondientes interfaces. En primer lugar, se incorporó al filtro existente por la palabra clave “mantenimiento” la búsqueda por los distintos tipos de mantenimiento definidos en la guía UIGreenmetric. Para ello, el código lee una tabla contenida en un archivo Word donde se especifican estos tipos, originalmente en inglés. Debido a que la búsqueda se realiza en castellano, se implementó una traducción previa de los términos antes de efectuar la búsqueda. Seguidamente, se implementó el informe “**Number of events related to environment and sustainability (6_8)**”, que permite al usuario seleccionar un rango de años y busca en los archivos de UBU Verde las actividades realizadas en ese periodo. Se consideran actividades sostenibles aquellas que se alinean con los Objetivos de Desarrollo Sostenible de las Naciones Unidas (<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>). Respecto al informe “**Total research funds dedicated to sustainability research (in US dollars) (6_4)**”, debido a su complejidad se decidió dividir la tarea en dos fases: desarrollo de la interfaz e implementación del código. Este informe recibe un archivo Excel con los datos de todos los fondos y los proyectos a los que se han destinado. A partir de este archivo, el código lee las diferentes columnas para calcular el dinero invertido en cada proyecto por año. Para determinar si un proyecto es sostenible, se realiza un análisis basado en el título del proyecto. El informe calcula el total invertido en todos los proyectos, así como el monto destinado exclusivamente a proyectos sostenibles para cada año. Posteriormente, busca el valor del dólar correspondiente a cada año y realiza la conversión de euros a dólares, calculando también un ratio entre ambas cantidades. Finalmente, se adjuntan dos tablas al informe: una con los resultados de esta conversión y otra con todos los cálculos detallados. No obstante, la implementación completa de este informe no se logró finalizar en este sprint, quedando pendiente la generación del Excel con los datos finales y la integración de las tablas al informe para sprints futuros.

Aunque inicialmente se estimó que la duración de este sprint sería de 15 días, debido a las tareas pendientes se extendió a un total de 45 días.

La Figura A.6 resume la implementación de filtrado avanzado y nuevos informes de sostenibilidad.

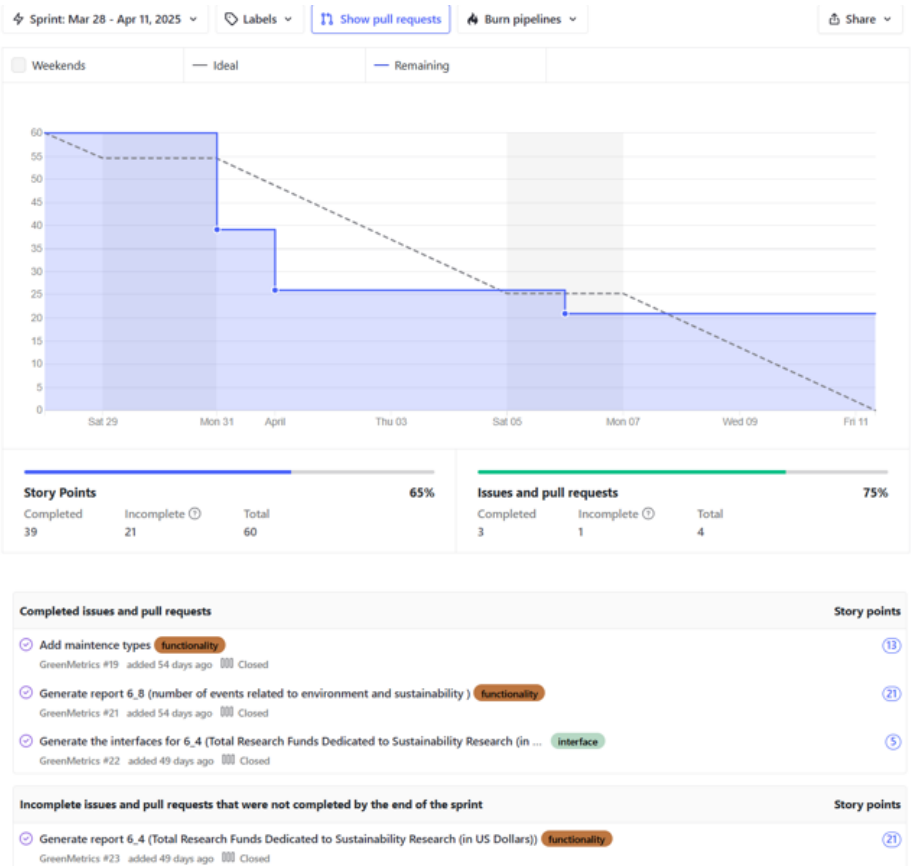


Figura A.6: Implementación de Filtrado Avanzado y Nuevos Informes de Sostenibilidad

Sprint(11/04/2025 – 25/04/2025)

– Sprint de Mejora de Interfaz, Tabla de Búsqueda e Implementación de Nuevos Informes –

Los objetivos principales de este sprint fueron mejorar la interfaz de la página principal, corregir errores relacionados con el registro e inicio de sesión, actualizar la tabla de consulta de datos y avanzar en la generación de los informes “**Total research funds dedicated to sustainability**”

research (in US dollars) (6_4)”, **“The ratio of sustainability courses to total courses/subjects (6_3)”** y **“Total number of courses or modules offered (6_2)”**. Se realizaron mejoras importantes en la página de búsqueda: se eliminó la columna con los nombres de los archivos de la tabla para una visualización más limpia y se añadió la posibilidad de filtrar los datos por año. También se implementó paginación, mostrando las filas de cien en cien, lo que mejora considerablemente la experiencia del usuario al consultar grandes volúmenes de información. Asimismo, se añadieron filtros interactivos y animaciones a los botones, aportando fluidez y dinamismo a la interfaz. En la página principal, se mejoró la visualización general: se actualizó el fondo, se rediseñó la barra de navegación y se agruparon los informes por categoría. Además, se añadieron animaciones y se resaltaron en negrita aquellos informes que ya están implementados. Como parte del trabajo de refactorización, se organizaron los archivos HTML y se creó una hoja de estilos CSS para evitar la duplicación de código y facilitar su mantenimiento. Durante este sprint se implementó el informe **“Total number of courses or modules offered (6_2)”**, el cual permite calcular el número total de asignaturas ofrecidas en grados y másteres, según el año seleccionado por el usuario. El sistema consulta automáticamente la base de datos de guías docentes y, si los datos no están disponibles, los descarga automáticamente sin necesidad de intervención manual. Por otro lado, se comenzó a escribir el código para los informes **“Total research funds dedicated to sustainability research (in US dollars) (6_4)”** y **“The ratio of sustainability courses to total courses/subjects (6_3)”**. Sin embargo, no se logró completar su desarrollo, por lo que su finalización queda pendiente para el próximo sprint.

Aunque inicialmente se estimó que la duración del sprint sería de 15 días, debido al volumen de trabajo, este se extendió a un total de 25 días.

La Figura A.7 presenta las mejoras realizadas en la interfaz de usuario y el avance en nuevos informes.

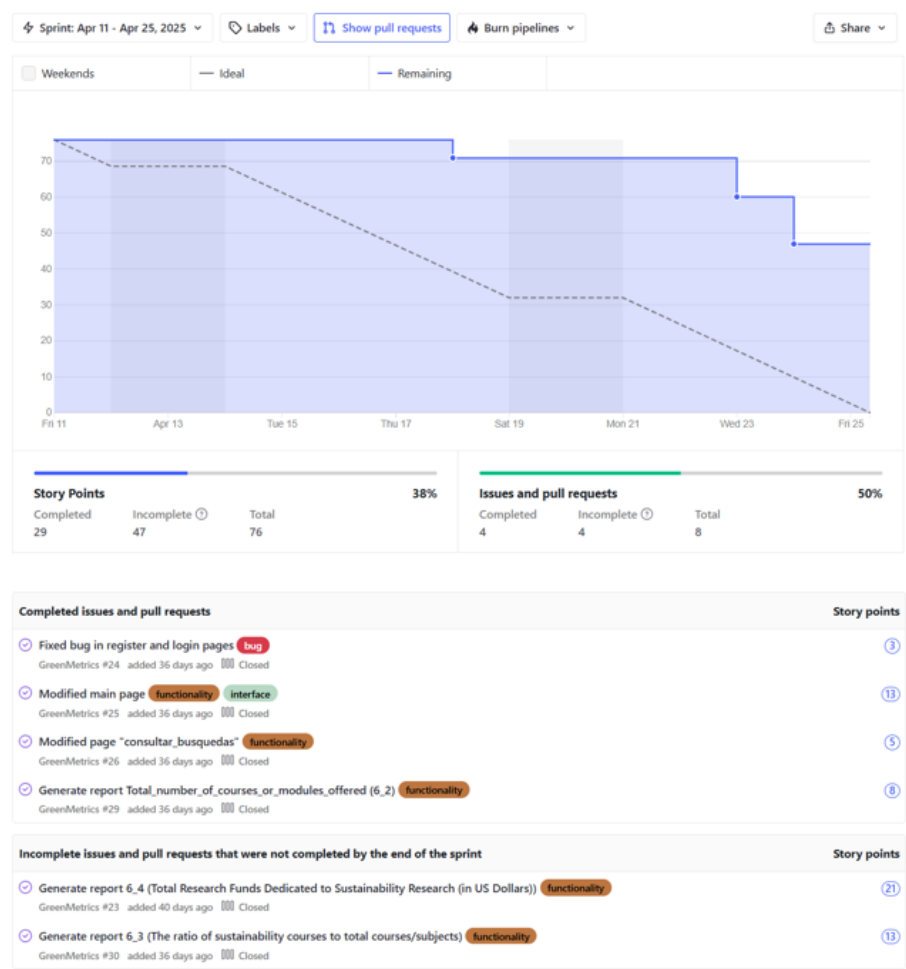


Figura A.7: Sprint de Mejora de Interfaz, Tabla de Búsqueda e Implementación de Nuevos Informes

Sprint(25/04/2025 – 09/05/2025)

– Sprint de Informes Académicos, Gestión de Roles y Documentación Técnica–

Los objetivos principales de este sprint fueron: mejorar el informe "Number of Courses on Environment and Sustainability (6_1)", finalizar el informe "Total research funds dedicated to sustainability research (in US dollars) (6_4)", generar el nuevo informe "The ratio of sustainability courses to total courses/subjects (6_3)", añadir un rol de administrador por defecto al crear la tabla de usuarios, y desarrollar una interfaz que permita designar y degradar administradores. Además, se

comenzó la redacción de la documentación técnica, incluyendo el manual de usuario y el README de GitHub. Se dio prioridad a completar el informe **"Total research funds dedicated to sustainability research (in US dollars) (6_4)"**, ya iniciado en sprints anteriores. En lugar de insertar capturas de pantalla de las tablas del archivo Excel, se optó por generar tablas directamente en el informe .docx, lo que facilita futuras ediciones sin necesidad de modificar manualmente el Excel y tomar nuevas capturas.

El informe **"Number of Courses on Environment and Sustainability (6_1)"** fue corregido tras detectarse un error durante la descarga de nuevos datos. Este error se debía a la inclusión de una nueva columna en la base de datos de las guías docentes, lo que requirió su reconstrucción. Además, se implementó una comprobación para evitar descargar datos ya existentes y se corrigió el almacenamiento erróneo del valor “ambos” en el campo “tipo de programa”, que ahora se guarda como “grado” o “máster” según corresponda.

Se implementó también una nueva interfaz de gestión de usuarios, permitiendo a los administradores modificar los roles. Solo los administradores pueden designar nuevos administradores, y únicamente el administrador creado por defecto puede degradar a otros administradores, garantizando así un mayor control y seguridad en el sistema de permisos. Además, los administradores tienen la capacidad de editar o eliminar filas de la base de datos de las guías docentes, mientras que todos los usuarios pueden personalizar las columnas visibles al consultar la tabla.

El nuevo informe **"The ratio of sustainability courses to total courses/subjects (6_3)"** fue desarrollado en este sprint. Este informe permite calcular el ratio de asignaturas sostenibles sobre el total, según el año seleccionado por el usuario. El sistema consulta automáticamente la base de datos de guías docentes para contar las asignaturas sostenibles en grados y másteres. En caso de que los datos necesarios no estén disponibles, el código los descarga automáticamente sin intervención del usuario.

Finalmente, se comenzó la redacción del manual de usuario y del README para GitHub. Sin embargo, no se logró completar la versión en inglés, por lo que su finalización se pospone para el siguiente sprint.

Aunque la duración estimada del sprint era de 15 días, debido a la carga de trabajo restante, se extendió a un total de 18 días.

En la Figura [A.8](#) se visualizan las tareas relacionadas con la gestión de roles y la documentación técnica.



Figura A.8: Sprint de Informes Académicos, Gestión de Roles y Documentación Técnica

Sprint(09/05/2025 – 23/05/2025)

–Sprint de Documentación y Organización del Proyecto –

Durante este sprint se abordaron diversas tareas clave orientadas a mejorar la calidad del proyecto, tanto a nivel técnico como de documentación.

Uno de los primeros pasos fue la integración de SonarCloud con el repositorio, lo cual permitió activar el análisis automático de la calidad del

código. Además, se añadió el badge correspondiente en el archivo README, que también fue redactado en dos idiomas: español e inglés.

Otro punto importante fue la decisión sobre el entorno de edición para la documentación técnica. Tras evaluar las opciones disponibles, se valoró si emplear LaTeX de forma local o utilizar la plataforma Overleaf, aunque no se tomó una decisión definitiva en este sprint. Paralelamente, se inició la redacción del manual de usuario utilizando la wiki de GitHub. Cada sección del manual se vinculó con su respectiva interfaz en la aplicación, y al igual que el README, el manual fue elaborado tanto en español como en inglés, lo que supuso un esfuerzo adicional de redacción y revisión.

En cuanto a la infraestructura para compartir el código y garantizar su portabilidad, se analizó la conveniencia de utilizar Docker frente a una máquina virtual. Finalmente, se optó por esta última opción debido a su mayor simplicidad y a que, dadas las múltiples interfaces de la aplicación, Docker podría introducir errores adicionales.

También se modificó la estructura de los directorios donde se almacenaban los informes generados, organizándolos de una forma más visual e intuitiva. Además, fue necesario actualizar las URL desde las cuales se descargaban las guías docentes. Esto se debió a una reestructuración en el sitio web de la Universidad de Burgos, que afectó tanto a las guías de grado como de máster. En consecuencia, se corrigieron los enlaces que alimentaban los informes titulados: "**Number of Courses on Environment and Sustainability (6_1)**", "**Total Number of Courses or Modules Offered (6_2)**" y "**The ratio of sustainability courses to total courses/subjects (6_3)**".

Otra tarea pendiente en este sprint fue la grabación de dos videotutoriales: uno breve, con una introducción general al uso de la aplicación, y otro más extenso, con una explicación detallada de cada funcionalidad. Se logró completar únicamente el videotutorial corto, el cual fue grabado en español e inglés, mientras que el vídeo detallado quedó pendiente para el siguiente sprint. De igual manera, no se pudo avanzar en la redacción de las primeras secciones de la memoria ni del anexo, debido al elevado volumen de trabajo acumulado.

Aunque inicialmente se había estimado una duración de 15 días para este sprint, la carga de tareas y la necesidad de completar ciertos elementos provocaron una extensión del mismo hasta los 17 días.

En la Figura A.9 se visualizan las tareas abordadas durante este sprint.

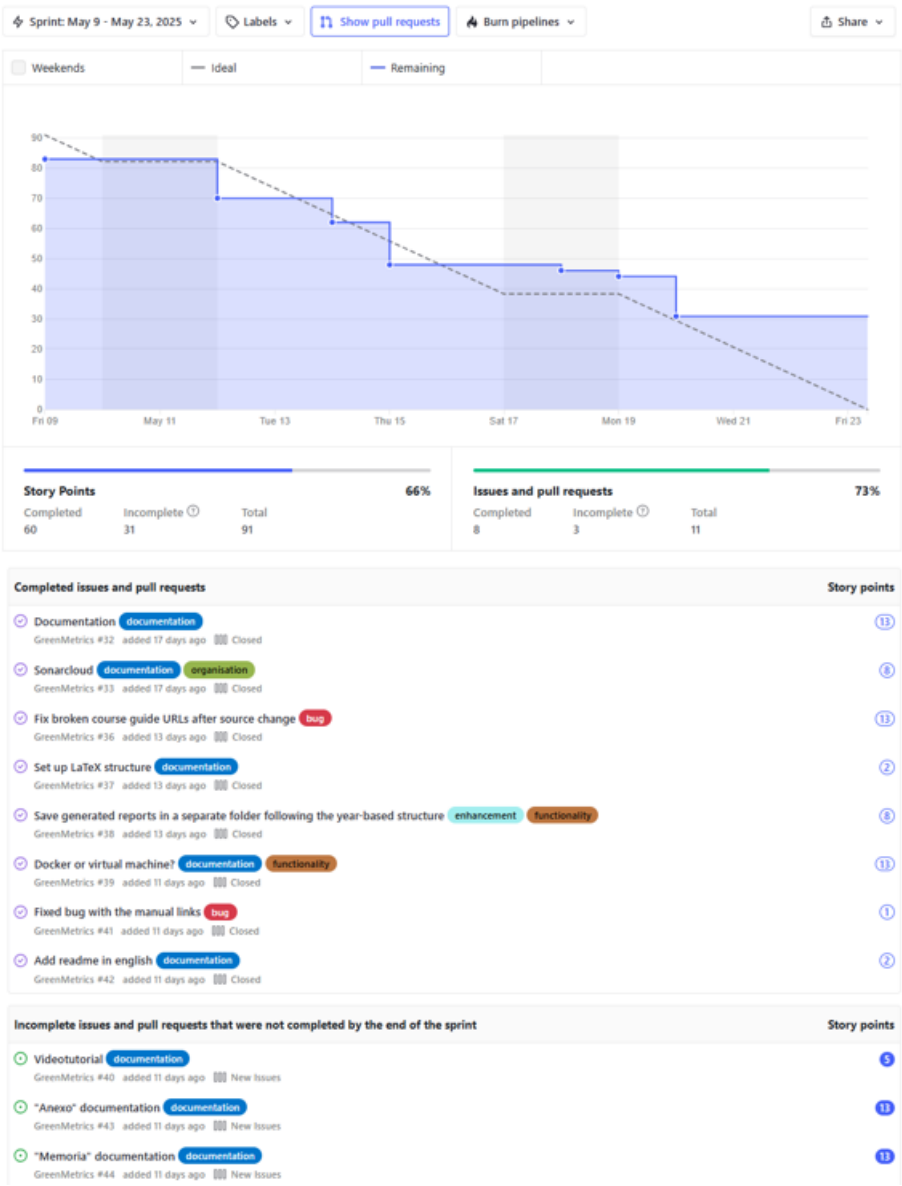


Figura A.9: Sprint de Documentación y Organización del Proyecto

Sprint(23/05/2025 – 06/06/2025)

Resumen

A continuación se muestra una Tabla A.2 donde aparece el número de tareas de las distintas categorías.

Categoría	Número de issues
Bug	5
Documentation	15
Enhacement	1
Functionality	27
Interface	11
Organisation	3

Tabla A.2: Horas dedicadas al proyecto

En el siguiente gráfico [A.10](#) se puede observar como la gran mayor parte de las tareas realizadas estaban destinadas a mejorar la funcionalidad de la aplicación y a la realización de la documentación.

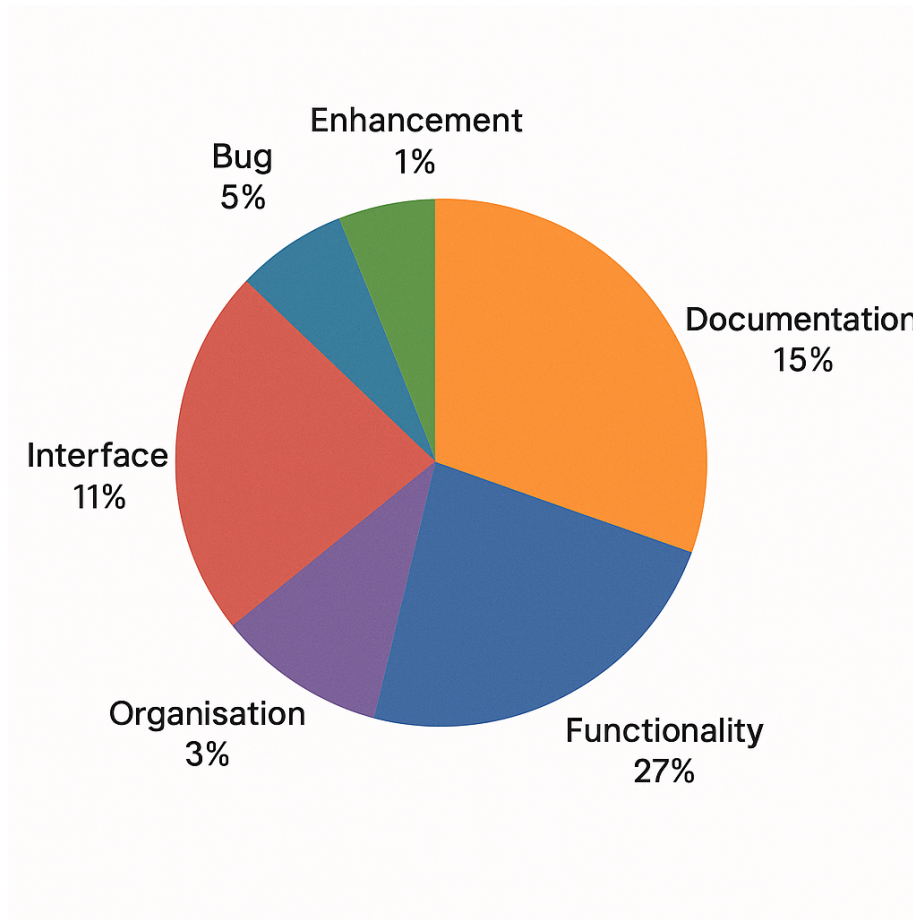


Figura A.10: Distribución de Issues por Categoría

A.3. Estudio de viabilidad

En este apartado se va a calcular el impacto económico de lanzar la aplicación al mercado.

Viabilidad económica

Costes

- **Coste de personal** En esta sección se va a calcular el gasto destinado a tener un empleado contratado. Se calcula que se han empleado unas 800 horas de trabajo repartidas en 6 meses. Esto supone, de manera aproximada, unas 31 horas semanales. El salario de un recién titulado de ingeniería informática, oscila en torno a los 1900€ mensuales; traba-

jando a jornada completa, que son generalmente 40 horas semanales. Como en este caso, solo se trabajan 31 horas a la semana:

$$\frac{1900 \text{ €}}{40 \text{ horas/mes}} \times 31 \text{ horas} = 1472,5 \text{ € al mes}$$

Este será el salario bruto del empleado, ahora calculamos el salario neto que recibirá y los impuestos que la empresa debe pagar por él (los impuestos se han sacado del siguiente [3]enlace).

La Figura A.11 y Figura A.12 se muestra el esquema de aportaciones a la Seguridad Social aplicables al cálculo del coste laboral.

TIPOS DE COTIZACIÓN (%)				
CONTINGENCIAS	EMPRESA	TRABAJADORES	TOTAL	Accidentes de Trabajo y Enfermedades Profesionales
Comunes	23,60	4,70	28,30	
Horas Extraordinarias Fuerza Mayor	12,00	2,00	14,00	Tarifa Primas establecida en la disposición adicional cuarta Ley 42/2006, de 28 de diciembre, PGE 2007, en la redacción dada por la Disposición Final Quinta del RDL 28/2018 de 28 de diciembre (BOE del 29) siendo las primas resultantes a cargo exclusivo de la empresa
Resto Horas Extraordinarias	23,60	4,70	28,30	
Mecanismo Equidad Intergeneracional (MEI)	0,67	0,13	0,8	

Figura A.11: Régimen general de la Seguridad Social I.

DESEMPLEO				EMPRESA	TRABAJADORES	TOTAL
Tipo General: Contratación indefinida, incluidos los contratos indefinidos a tiempo parcial y fijos discontinuos, contratación de duración determinada en las modalidades de contratos de formación en alternancia, para la formación y aprendizaje, formativo para la obtención de la práctica profesional adecuada al nivel de estudios, de relevo, interinidad y contratos realizados con trabajadores que tengan reconocido un grado de discapacidad no inferior al 33%				5,50	1,55	7,05
Contrato duración determinada a tiempo completo o a tiempo parcial				6,70	1,60	8,30

	EMPRESA	TRABAJADORES	TOTAL
E.Q.G.A.S.A	0,20		0,20

	EMPRESA	TRABAJADORES	TOTAL
FORMACIÓN PROFESIONAL	0,60	0,10	0,70

Figura A.12: Régimen general de la Seguridad Social II.

De las imágenes sacamos la siguiente informacion

- **Contingencias** = $23,60 \% = 0,236$
- **Mecanismo Equidad Intergeneracional (MEI)** = $0,67 \% = 0,0067$

- **Desempleo** = $5,50\% = 0,055$
- **FOGASA** = $0,20\% = 0,002$
- **Formación profesional** = $0,60\% = 0,006$

Por lo que:

$$1472,5 \text{ € al mes} \times (1 + (0,236 + 0,0067 + 0,055 + 0,002 + 0,006)) = 1922,64 \text{ € al mes}$$

También tendremos en cuenta el sueldo de los dos profesores. El salario de cada uno es de 30€ la hora, ya que cuenta con experiencia. Se les contrata por una hora a la semana; teniendo en cuenta que cada mes tiene unas cuatro semanas:

$$\begin{aligned} 1 \text{ hora a la semana} \times 30 \text{ € la hora} \times 4 \text{ semanas al mes} &= \\ &= 120 \text{ € al mes por profesor} \end{aligned}$$

Este será el salario bruto de los profesores, ahora calculamos el salario neto que recibirán y los impuestos que la empresa debe pagar:

$$\begin{aligned} (120 \text{ € al mes} \times 2 \text{ profesores}) \times (1 + (0,236 + 0,0067 + 0,055 + 0,002 + 0,006)) &= \\ &= 313,37 \text{ € al mes} \end{aligned}$$

Como el proyecto a durado un total de seis meses, el coste para la empresa es de :

$$\begin{aligned} (1922,64 \text{ € sueldo del empleado} + 313,37 \text{ € sueldo de los profesores}) \times 6 \text{ meses proyecto} &= \\ &= 13416,06 \text{ €} \end{aligned}$$

- **Hardware** El recursos hardware empleado ha sido un ordenador, cuyo coste fue de 800€. Por lo que los costes han sido 80€.

$$\left(\frac{800 \text{ €}}{5 \text{ años}} \right) \times \frac{6 \text{ meses de proyecto}}{12 \text{ meses}} = 80 \text{ €}$$

- **Software** Todas las herramientas software utilizadas han sido gratuitas, por lo que no han tenido coste alguno.
- **Costes indirectos** Para calcular los costes, es necesario tener en cuenta los costes indirectos, como la tarifa de internet. En este caso, la tarifa es de 28€/mes. Como la duración ha sido de 6 meses, el total de los costes indirectos es 168€.

Beneficios

Dado que el proyecto está destinado al ámbito académico, concretamente a investigadores de la Universidad de Burgos, su uso no conlleva fines comerciales ni la obtención de beneficios.

Tipo coste	Total
Personal	13416,06€
Hardware	80€
Software	0€
Indirectos	168€
Total	13664,06€

Tabla A.3: Costes- totales

Viabilidad legal

En este apartado se van a estudiar las diferentes licencias de las herramientas empleadas.

La Tabla [A.4](#) presenta una comparación entre las licencias de software utilizadas.

Herramienta	Versión	Licencia
Visual Studio Code	1.100.2	MIT
Flask	3.1.0	BSD
GitHub	2.65.0	GNU
Flask-SQLAlchemy	3.1.1	MIT
LM studio	0.3.9	MIT
SQLAlchemy	2.0.36	MIT
requests	2.32.3	Apache 2.0
selenium	4.29.0	Apache 2.0
pandas	2.2.3	BSD
numpy	1.26.2	BSD
matplotlib	3.8.2	PSF / BSD
google-auth	2.38.0	Apache 2.0
Werkzeug	3.1.3	BSD
WTForms	3.2.1	BSD
xlwings	0.33.12	BSD

Tabla A.4: Licencias de herramientas y librerías

A continuación compararemos las distintas licencias en la Tabla A.5.

Característica	MIT	BSD	Apache 2.0	GPL
Permisividad	Alta	Alta	Media	Baja
Uso comercial permitido	Sí	Sí	Sí	Sí
Compartir modificaciones	Opcional	Opcional	Recomendado	Obligatorio
Protección de patentes	No	No	Sí	No
Requisitos de atribución	Sí	Sí	Sí	Sí
Compatibilidad con otros proyectos	Alta	Alta	Media	Baja

Tabla A.5: Comparación de licencias

Como se puede observar, tanto la licencia MIT como la licencia BSD comparten características similares en cuanto a permisividad y compatibilidad. La licencia MIT se caracteriza por su simplicidad, alta permisividad y gran compatibilidad con otros proyectos. Permite que cualquier persona modifique el código e integrarlo incluso en software de código cerrado, sin necesidad de compartir públicamente las modificaciones.

Por otro lado, la licencia BSD, especialmente en su versión de 3 cláusulas, ofrece una flexibilidad comparable a la de MIT, pero impone una restricción adicional: prohíbe el uso del nombre del autor o de la organización para promocionar productos derivados sin autorización. Esto proporciona un control adicional sobre el reconocimiento del autor, mientras se mantiene la libertad de uso y modificación del código.

Dado que la intención de este trabajo es fomentar la accesibilidad y reutilización del código sin restricciones, se opta por la **licencia MIT** por su simplicidad y amplia compatibilidad con otros proyectos.

Apéndice B

Especificación de Requisitos

B.1. Introducción

En este apartado se especificarán los requisitos que debe tener la aplicación.

B.2. Objetivos generales

Este proyecto tiene como objetivos generales los siguientes:

- Crear una aplicación que automatice la creación de informes.
- Para cada informe solicitar los datos que dicho informe necesite, como por ejemplo el curso académico, y si dicho informe necesita realizar búsquedas en internet que lo haga sin la intervención del usuario.
- Generar una base de datos donde se almacenen las guías docentes que se han ido descargando para facilitar la consulta.

B.3. Catálogo de requisitos

Los requisitos se organizan en funcionales y no funcionales, garantizando la correcta implementación y rendimiento del sistema.

B.3.1 Requisitos Funcionales

Los requisitos funcionales establecen las capacidades que el sistema debe cumplir para su correcto funcionamiento.

- **RF1- Registro de usuario** La aplicación debe permitir a los usuarios crear una cuenta y que sus datos sean almacenados en la base de datos.
- RF2- Registro de login** La aplicación debe permitir a los usuarios acceder a las distintas funcionalidades, dependiendo de su rol.

RF3- Procesamiento de guías docentes

- RF3.1- Extraer y analizar el contenido de archivos PDF con información sobre asignaturas universitarias.
- RF3.2- Identificar competencias en sostenibilidad y relacionarlas con los ODS . Almacenando un "Sí." "No." en función de si son sostenibles o no.
- RF3.3- Generar informes de análisis en Excel para consulta académica.

RF4- Automatización con modelos de IA

- RF4.1- Construcción de prompts dinámicos para consulta en una API de LLMs.
- RF4.2- Procesamiento de texto y detección de información clave sobre sostenibilidad.
- RF4.3- Filtrado y clasificación de respuestas generadas por la IA.

RF5- Gestión de base de datos

- RF5.1- Almacenar información en una base de datos SQL mediante Flask-SQLAlchemy .
- RF5.2- Comprobar existencia de registros previos para evitar duplicaciones.
- RF5.3- Actualizar asignaturas cuando se detecten cambios en su contenido.
- RF5.4- Editar las filas: que se podrán actualizar o eliminar.
- RF5.5- Mostrar la base de datos, donde el usuario podrá consultar las guías docentes que hayan sido descargadas.

RF6- Configuración de la aplicación

- RF6.1- El usuario podrá configurar el idioma, tamaño de la letra, y escoger la opción de daltonismo.
- RF6.2- El usuario podrá modificar sus datos como nombre de usuario, correo electrónico y contraseña. Además de poder cambiar de rol a otros usuarios, en caso de ser administrador.

RF7- Ayuda de la aplicación El usuario debe poder obtener ayuda en cada una de las funcionalidades de la aplicación.

RF8- Informe 1__19 Recopilación y procesamiento automatizado de contratos de mantenimiento de edificios universitarios mediante web scraping y procesamiento de texto.

RF9- Informe 6__1 Extracción de información sobre cursos y programas académicos relacionados con sostenibilidad mediante técnicas de análisis de planes de estudio y web scraping.

RF10- Informe 6__2 Cálculo del número total de asignaturas ofrecidas en la universidad, generando evidencias estructuradas en la base de datos.

RF11- Informe 6__3 Análisis de la proporción de cursos de sostenibilidad respecto al total de cursos, permitiendo evaluar el compromiso académico con la sostenibilidad.

RF12- Informe 6__4 Procesamiento de información financiera para calcular la inversión en investigación sobre sostenibilidad, incluyendo métricas de comparación por años.

RF13- Informe 6__7 Recopilación automatizada del número de publicaciones científicas en sostenibilidad de la Universidad de Burgos mediante Google Scholar.

RF14- Informe 6__8 Identificación del número de eventos relacionados con sostenibilidad mediante la clasificación de actividades de UBU Verde con modelos de lenguaje.

B.3.2 Requisitos No Funcionales

Los requisitos no funcionales establecen condiciones que afectan el rendimiento y la experiencia de usuario.

■ RNF1- Eficiencia y rendimiento

- RNF1.1- Optimización del procesamiento de PDFs para minimizar tiempos de ejecución.

- RNF1.2- Manejo eficiente de grandes volúmenes de archivos sin afectar estabilidad.

RNF2- Seguridad y privacidad

- RNF2.1- Protección de credenciales API mediante variables de entorno .
- RNF2.2- Implementación de controles de integridad en el almacenamiento de datos.

RNF3- Mantenibilidad y escalabilidad

- RNF3.1- Separación modular en distintos componentes: extracción, procesamiento y almacenamiento.
- RNF3.2- Adaptabilidad para integrar nuevos modelos de IA o fuentes externas.

RNF3- Compatibilidad y accesibilidad

- RNF3.1-Funcionalidad multiplataforma (Linux, Windows, macOS).
- RNF3.2- Uso de formatos estándar como PDF y Excel para facilitar integración.

RNF4- Internacionalización la aplicación deberá estar preparada para soportar varios idiomas (castellano e inglés).

B.4. Especificación de requisitos

CU-1	Registro de usuario
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-1
Descripción	Permite a los usuarios registrarse, quedando guardados los datos en la base de datos.
Precondición	Existe la base de datos
Acciones	<ol style="list-style-type: none"> 1. El usuario mete los datos correspondientes: nombre de usuario, correo y contraseña (que debe tener un mínimo de ocho caracteres, donde al menos haya una mayúscula, una minúscula y un número.
Postcondición	<p>Que todos los datos introducidos sean válidos.</p> <ul style="list-style-type: none"> ■ Se verifican que el nombre de usuario, no este ya en uso en la base de datos. ■ Se verifica que el correo electrónico sea válido, y no este ya en la base de datos. ■ Como la contraseña se introduce dos veces, se verifica que en ambos casos sea la misma.
Excepciones	<ul style="list-style-type: none"> ■ Nombre de usuario en uso. ■ No se ha introducido un correo válido o ya esta en uso. ■ La contraseña no cumple con los requisitos de seguridad o no coinciden.
Importancia	Alta

Tabla B.1: CU-1 Registro de usuario.

CU-2	Registro de login
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-2
Descripción	Permite a los usuarios iniciar sesión.
Precondición	Existe la base de datos y los datos del usuario en dicha base
Acciones	<ol style="list-style-type: none"> 1. El usuario mete los datos correspondientes: nombre de usuario, correo y contraseña.
Postcondición	<p>Que todos los datos introducidos sean correctos.</p> <ul style="list-style-type: none"> ■ Se verifican que el nombre de usuario/correo y contraseña, coincidan con la que esta almacenada en la base de datos.
Excepciones	<ul style="list-style-type: none"> ■ Nombre de usuario/correo incorrecto. ■ La contraseña incorrecta.
Importancia	Alta

Tabla B.2: CU-2 Registro de login.

CU-3	Procesamiento de guías docentes
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-3, RF-3.1, RF-3.2, RF-3.3
Descripción	Permite a los usuarios descargarse guías docentes.
Precondición	No existen esas guías docentes en la base de datos
Acciones	<ol style="list-style-type: none"> 1. El usuario introduce el curso académico del que desea descargar las guías docentes, y si quiere descargar las de grado, master o ambas. 2. Comienza la descarga de dichas guías. 3. Se lee el contenido del PDF de cada guía docente para extraer la información necesaria (por ejemplo, código de la asignatura, título). 4. En base a las competencias de cada guía identificar si alguna es sostenible. 5. Generar un excel y actualizar la base de datos con las guías que han sido descargadas.
Postcondición	Que todas las guías se hayan descargado.
Excepciones	-
Importancia	Media

Tabla B.3: CU-3 Procesamiento de guías docentes.

CU-4	Automatización con modelos de IA
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-4, RF-4.1, RF-4.2, RF-4.3
Descripción	Permite a los usuarios utilizar un LLM local.
Precondición	Tener iniciado el LLM.
Acciones	<ol style="list-style-type: none"> 1. El usuario selecciona una guía docente para su análisis. 2. El sistema construye un prompt dinámico con la información relevante. 3. Se envía el prompt al LLM local para su procesamiento. 4. El LLM detecta y extrae información clave sobre sostenibilidad. 5. Se realiza un filtrado y clasificación de las respuestas generadas. 6. Los datos procesados se almacenan en la base de datos y se generan evidencias.
Postcondición	La información analizada queda almacenada y disponible en la base de datos.
Excepciones	Error en el procesamiento del LLM o fallo en la comunicación con la API.
Importancia	Alta

Tabla B.4: CU-4 Automatización con modelos de IA.

CU-5	Gestión de base de datos
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-5, RF-5.1, RF-5.2, RF-5.3
Descripción	Actualiza la base de datos
Precondición	No existen esos mismos datos en la base de datos
Acciones	<ol style="list-style-type: none"> 1. Se comprueba que esos datos que se estan descargando no existan ya en la base de datos, para evitar duplicados. 2. Se actualiza la base de datos.
Postcondición	-
Excepciones	-
Importancia	Media

Tabla B.5: CU-5 Gestión de base de datos

CU-6	Editar/eliminar filas
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-5.4
Descripción	Permite a los administradores hacer modificaciones en la base de datos.
Precondición	Ser administrador y que exista la base de datos
Acciones	<ol style="list-style-type: none"> 1. El usuario escoge la acción (editar o eliminar) de la fila que quiera. 2. En caso de editar, se le mostrará una ventana donde podrá modificar los datos para esa guía docente en concreto. 3. En caso de seleccionar eliminar, se le pedirá una confirmación ya que dicha acción no se puede deshacer.
Postcondición	La información de la base de datos se actualiza.
Excepciones	-
Importancia	Media

Tabla B.6: CU-6 Editar/eliminar filas

CU-7	Mostrar base de datos
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-5.5
Descripción	Permite a los usuarios ver los datos almacenados en la base de datos.
Precondición	Que exista la base de datos.
Acciones	<ol style="list-style-type: none"> 1. El usuario selecciona el año del que desea visualizar los datos. 2. Selecciona si lo desea, algún filtro para la búsqueda. 3. Selecciona, que columnas quiere que se les muestren.
Postcondición	Se muestra la información de la base de datos.
Excepciones	-
Importancia	Media

Tabla B.7: CU-7 Mostrar base de datos

CU-8	Configuración
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-6.1
Descripción	Permite a los usuarios cambiar la configuración de la aplicación.
Precondición	-
Acciones	<ol style="list-style-type: none"> 1. El usuario escoge el idioma, tamaño de letra y si quiere la opción de daltonismo.
Postcondición	La aplicación se actualiza en base a la configuración seleccionada.
Excepciones	-
Importancia	Media

Tabla B.8: CU-8 Configuración

CU-9	Configuración de perfil
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-6.2
Descripción	Permite a los usuarios sus datos.
Precondición	Haber iniciado sesión.
Acciones	<ol style="list-style-type: none"> 1. El usuario puede cambiar su nombre de usuario, correo electrónico y/o contraseña.
Postcondición	La base de datos se actualiza en base a los nuevos datos introducidos.
Excepciones	-
Importancia	Media

Tabla B.9: CU-9 Configuración perfil

CU-4	Ayuda de la aplicación
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-7
Descripción	Permite a los usuarios consultar el manual de ususario.
Precondición	Haber pulsado el botón del manual.
Acciones	<ol style="list-style-type: none"> 1. Se redirige al usuario a la parte del manual de la interfaz correspondiente a donde ha pulsado la ayuda del manual.
Postcondición	-
Excepciones	-
Importancia	Media

Tabla B.10: CU-10 Ayuda de manual

CU-10	Informe 1_19
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-8
Descripción	Genera la evidencia 1_19.
Precondición	Tener el fichero con los tipos de mantenimiento.
Acciones	<ol style="list-style-type: none"> 1. Se traduce el listado de tipos de mantenimineto de inglés a castellano. 2. Se hace web scraping en la página de licitaciones de la Universidad de Burgos. 3. Se extraen los datos para cada tipo de mantenimiento. 4. Se genera el informe.
Postcondición	Se genera la evidencia correspondiente
Excepciones	-
Importancia	Alta

Tabla B.11: CU-11 Informe 1_19

CU-12	Informe 6_1
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-9
Descripción	Genera la evidencia 6_1
Precondición	Tener descargadas las guías docentes y haber seleccionado un año.
Acciones	<ol style="list-style-type: none"> 1. Para el año introducido, se extrae la información de las guías docentes. 2. Se genera la evidencia.
Postcondición	Se ha generado la evidencia.
Excepciones	No están las guías docentes descargadas con anterioridad.
Importancia	Alta

Tabla B.12: CU-12 Informe 6_1

CU-13	Informe 6_2
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-10
Descripción	Genera la evidencia 6_2
Precondición	Haber introducido un año.
Acciones	<ol style="list-style-type: none"> 1. Se consulta la base de datos, contando el total de datos para dicho año.
Postcondición	Se ha generado la evidencia.
Excepciones	No hay datos en la base de datos para dicho año; y tienen que estar disponible tanto para master como para grado.
Importancia	Alta

Tabla B.13: CU-13 Informe 6_2

CU-14	Informe 6_3
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-11
Descripción	Genera la evidencia 6_3
Precondición	Haber seleccionado un año.
Acciones	<ol style="list-style-type: none"> 1. Se consulta la base de datos, contando el total de datos que son y no son sostenibles para dicho año. 2. Calcula el ratio.
Postcondición	Se ha generado la evidencia
Excepciones	No están disponibles los datos para grado y master.
Importancia	Alta

Tabla B.14: CU-14 Informe 6_3

CU-14	Informe 6_3
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-12
Descripción	Genera la evidencia 6_4
Precondición	Haber adjuntado un documento.
Acciones	<div>1. Procesa la información del excel adjuntado.</div> <div>2. Para cada proyecto en base al título determina si es o no sostenible usando un LLM local.</div> <div>3. Calcula el fondo invertido por año de duración del proyecto.</div> <div>4. Calcula el fondo total destinado a proyectos sostenibles y proyectos no sostenibles.</div> <div>5. Genera un fichero excel con los cálculos, y adjunta la tabla en el informe.</div>
Postcondición	Se ha generado la evidencia
Excepciones	No se ha adjuntado ningún fichero con un formato válido.
Importancia	Alta

Tabla B.15: CU-15 Informe 6_4

CU-16	Informe 6_7
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-13
Descripción	Genera la evidencia 6_7
Precondición	Haber seleccionado un año.
Acciones	<ol style="list-style-type: none"> 1. Extrae el número de publicaciones en GoogleScholar para un año dado.
Postcondición	Se ha generado la evidencia
Excepciones	No se ha seleccionado un año.
Importancia	Alta

Tabla B.16: CU-16 Informe 6_7

CU-17	Informe 6_8
Versión	1.0
Autor	Lorena Bueno Porras
Requisitos asociados	RF-14
Descripción	Identifica el número de eventos sostenibles leyendo la información de un fichero word.
Precondición	Haber seleccionado un año.
Acciones	<ol style="list-style-type: none"> 1. Busca los ficheros que estén dentro del rango de años introducido por el usuario. 2. Va extrayendo de cada uno de ellos, cuantas actividades sostenibles se han realizado. 3. Genera el informe.
Postcondición	Se ha generado la evidencia
Excepciones	No hay ningún fichero en la carpeta <code>/src/generar_informe/UBU_{verde}_informes</code>
Importancia	Alta

Tabla B.17: CU-17 Informe 6_8

Apéndice C

Especificación de diseño

C.1. Introducción

En este apartado se explica cómo se han organizado y diseñado las diferentes partes de la aplicación.

C.2. Diseño de datos

La base de datos cuenta con dos tablas, las cuales no se relacionan entre sí:

- database: es la tabla donde se almacenan los datos de los usuarios (nombre de usuario, correo electrónico, contraseña y rol).
- busqueda: es la tabla donde se almacenan los datos de las guías docentes que se han descargado. Se guardan datos como :
 - Nombre del fichero.
 - Curso académico.
 - Código de la asignatura.
 - si es grado o master.
 - Modalidad
 - Si es sostenible o no.

C.3. Diseño arquitectónico

En este apartado se explicarán los diferentes patrones de diseño que se han utilizado para que la aplicación tenga un buen diseño software y esto ayude a la mantenibilidad de esta.

Modelo-Vista-Controlador (MVC)

El Modelo-Vista-Controlador (MVC) es un patrón de arquitectura de software que se utiliza para organizar y estructurar aplicaciones de manera modular, separando la lógica en tres componentes principales:

- **Modelo:** maneja los datos y la lógica de negocio.
- **Vista:** es la interfaz de usuario.
- **Controlador:** procesa las acciones del usuario y decide como actualizar el modelo y la vista.

Cada componente abarca:

- **Modelo:** Está compuesto por los módulos de procesamiento de guías docentes, almacenamiento de datos y análisis de sostenibilidad. En estos módulos se realizan las operaciones necesarias para extraer datos de archivos PDF, procesarlos con modelos de lenguaje grande (LLMs) y almacenarlos en la base de datos SQL mediante **Flask-SQLAlchemy**. También incluye la gestión de informes en formato Excel y la clasificación de asignaturas según criterios de sostenibilidad.
- **Vista:** Es la interfaz de usuario donde los usuarios pueden consultar informes sobre sostenibilidad en programas académicos y visualizar datos almacenados en la base de datos. Se ha diseñado en **HTML, CSS y JavaScript**, asegurando compatibilidad multiplataforma. La interfaz permite interacción con los datos descargados, mostrando tablas y además de opciones de configuración personalizadas como idioma y accesibilidad.
- **Controlador:** Es el módulo encargado de gestionar las comunicaciones entre la vista y el modelo, desarrollado en **Python con Flask**. Recibe las solicitudes de los usuarios, procesa las guías docentes con técnicas de procesamiento de lenguaje natural, consulta la base de datos y envía respuestas en formato JSON a la interfaz. También controla el

flujo de trabajo de los informes generados por el sistema, asegurando que los datos sean clasificados y presentados correctamente.

Las ventajas de utilizar este modelo es que se facilita la escalabilidad y el mantenimiento del código; además de permitir la reutilización de componentes pudiendo modificar la interfaz sin afectar a la lógica del código.

Se usa para que el procesamiento con LLMs y el acceso a datos sean manejados mediante una única entrada de control.

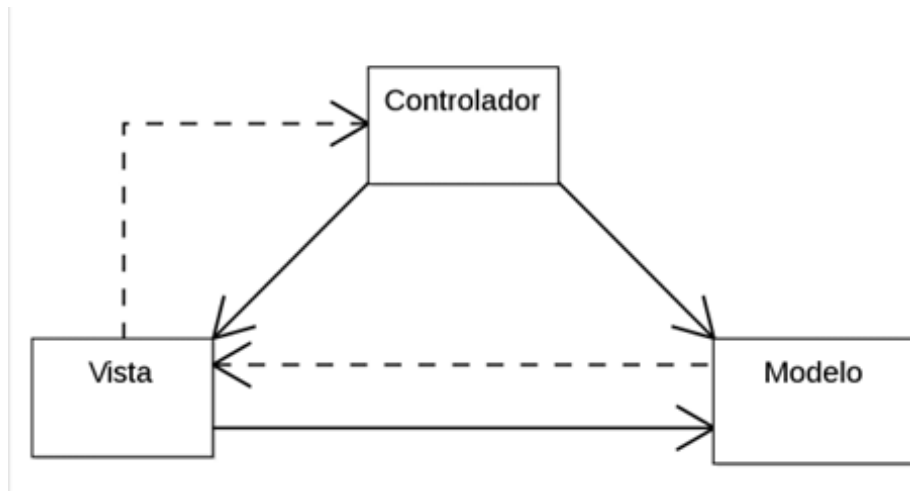


Figura C.1: Patrón MVC [5]

Patrón fachada

La aplicación implementa un módulo de control centralizado que actúa como interfaz única para gestionar las interacciones entre los distintos componentes del sistema. Este módulo simplifica el acceso a las funcionalidades, reduciendo la dependencia entre los subsistemas y permitiendo una mejor organización del código.

El funcionamiento del patrón fachada es el siguiente: Cuando un usuario solicita la generación de un informe sobre sostenibilidad, la fachada dirige la petición a los módulos internos de procesamiento de datos, base de datos y generación de informes. De esta manera, el usuario solo interactúa con una única entrada, mientras que el sistema administra internamente las operaciones necesarias para obtener los resultados.

Algunas de las ventajas de usar este patrón son que se simplifica la interacción entre módulos, evitando dependencias directas. También se

facilita la escalabilidad y se reduce la complejidad del código, mejorando su mantenimiento y optimización.

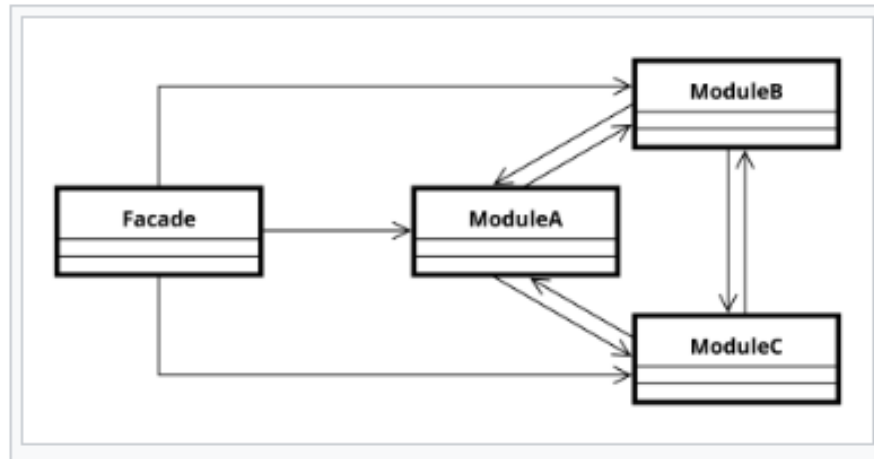


Figura C.2: Patrón fachada [6]

C.4. Diseño de interfaces

Para el desarrollo de las interfaces, se ha optado por no utilizar plantillas predefinidas. Esta decisión se tomó con el objetivo de adquirir experiencia práctica en HTML, CSS y JavaScript, lenguajes que no fueron abordados durante la formación académica. Además, este enfoque ha permitido profundizar en la comprensión de la interacción entre los distintos componentes de una interfaz, optimizando el diseño y la comunicación entre sus elementos

C.5. Diseño procedimental

En este apartado se muestra la ejecución de la aplicación mediante varios diagramas de secuencia.

Registro y Login

El primer diagrama de secuencia representa los pasos que sigue el programa en el momento en el que el usuario se registra y después accede a la aplicación con su cuenta.

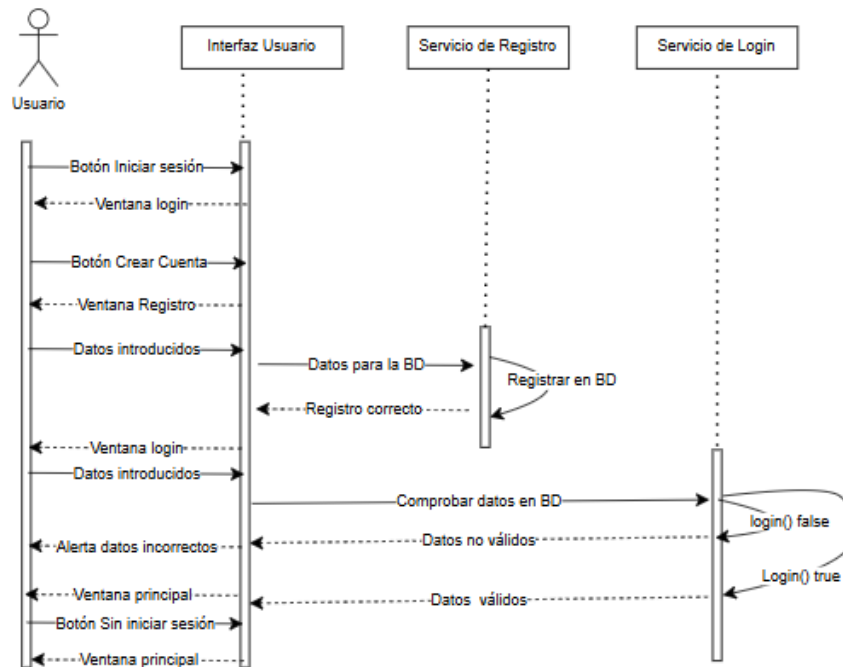


Figura C.3: Diagrama de secuencia de Registro y Login

Evidencias

A continuación se muestran los diagramas de secuencia para las distintas evidencias:

Evidencia Annual Operation Maintenance Percentage (1_19)

Aquí podemos observar los pasos al ejecutar la evidencia 1_19.

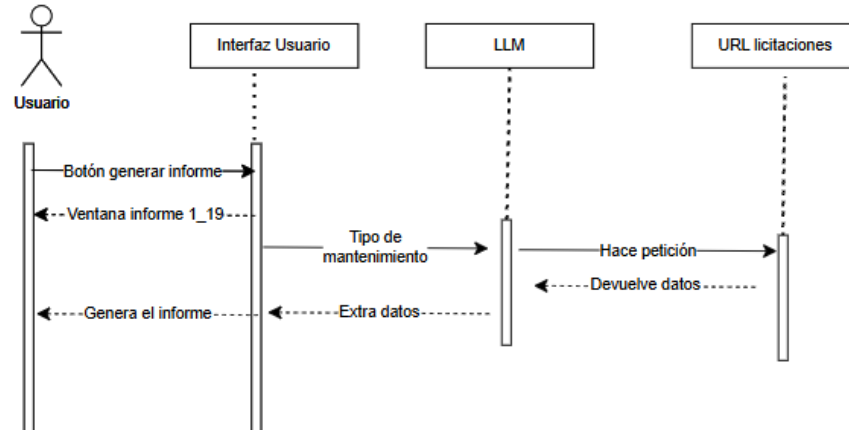


Figura C.4: Diagrama de secuencia evidencia 1_19

Number of Courses on Environment and Sustainability (6_1)

Aquí podemos observar los pasos al ejecutar la evidencia 6_1, tanto para la consulta de datos, como para la descarga de datos nuevos y la realización del informe.

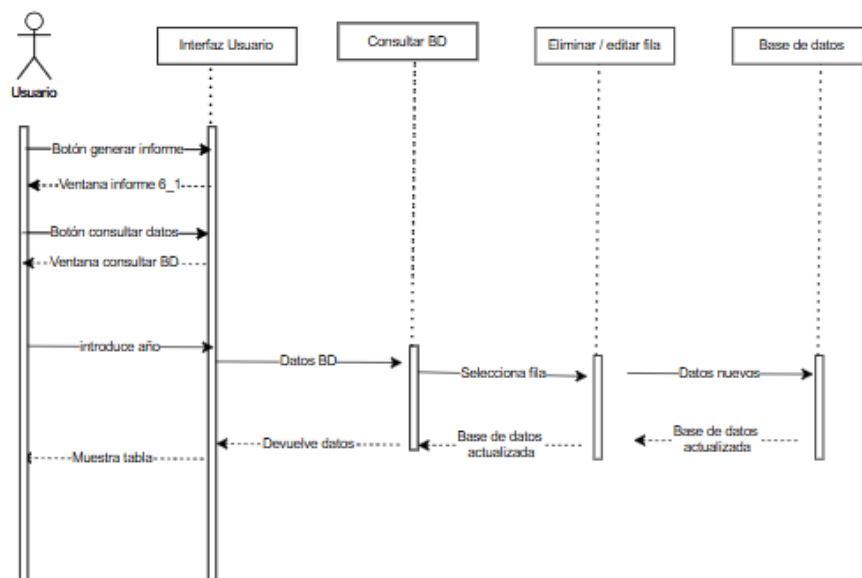


Figura C.5: Diagrama de secuencia de la evidencia 6_1 para consultar datos

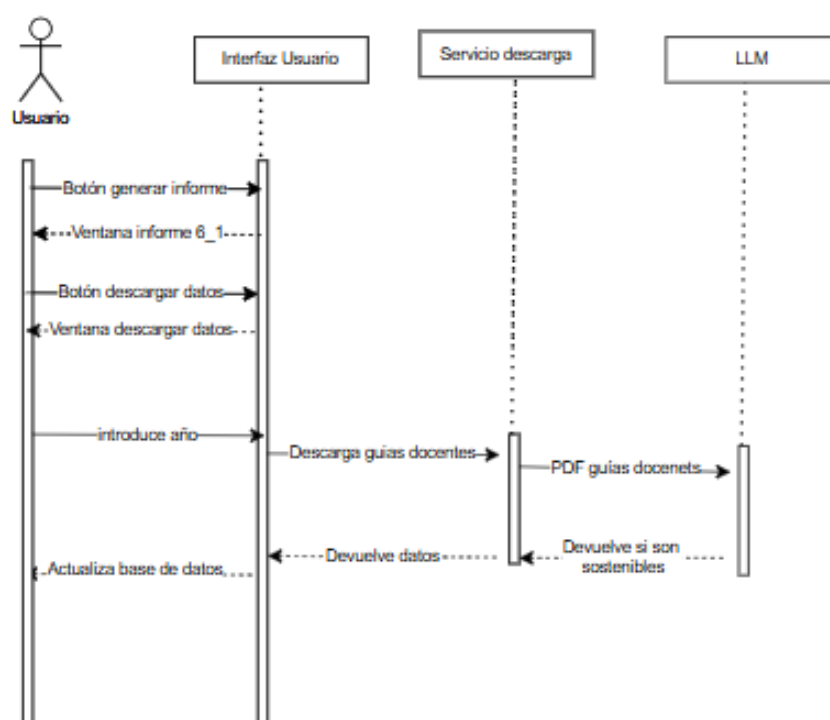


Figura C.6: Diagrama de secuencia de la evidencia 6_1 para descargar datos

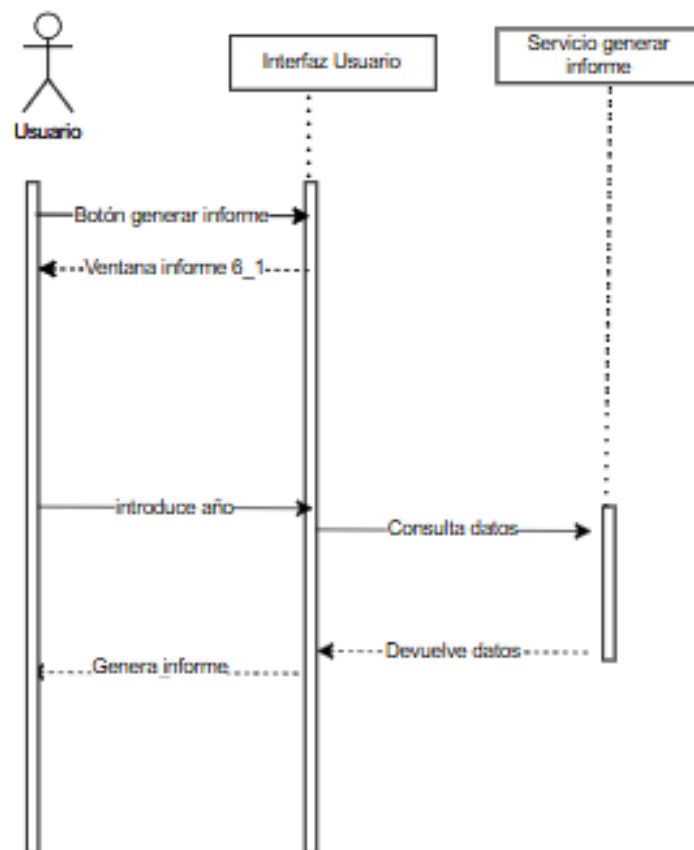


Figura C.7: Diagrama de secuencia de la evidencia 6_1 para generar el informe

Total number of courses or modules offered (6_2) y Ratio of sustainability courses to total courses/subjects (6_3)

Aquí podemos observar los pasos al ejecutar la evidencia 6_1 y 6_3.

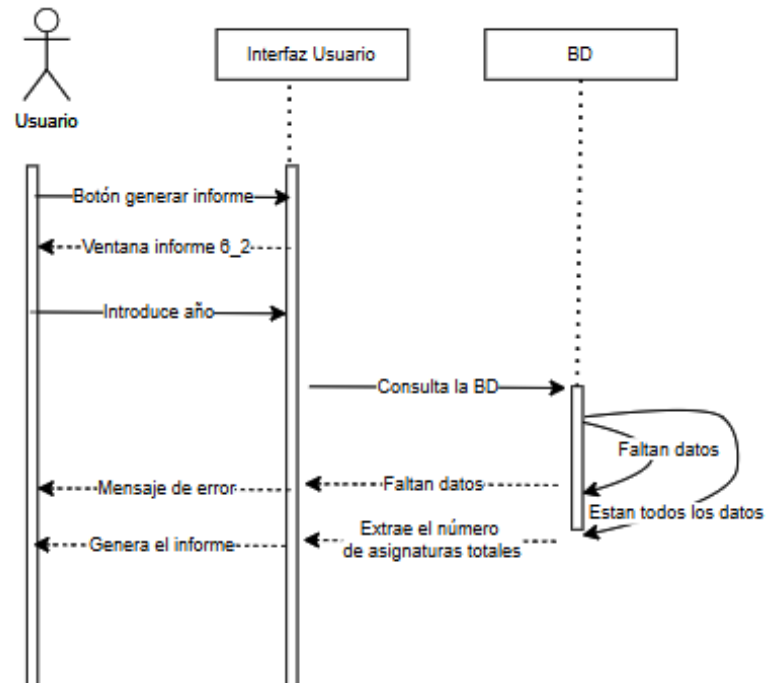


Figura C.8: Diagrama de secuencia de las evidencias 6_2 y 6_3

Total Research Funds Dedicated to Sustainability Research (in US Dollars) (6_4)

Aquí podemos observar los pasos al ejecutar la evidencia 6_4.

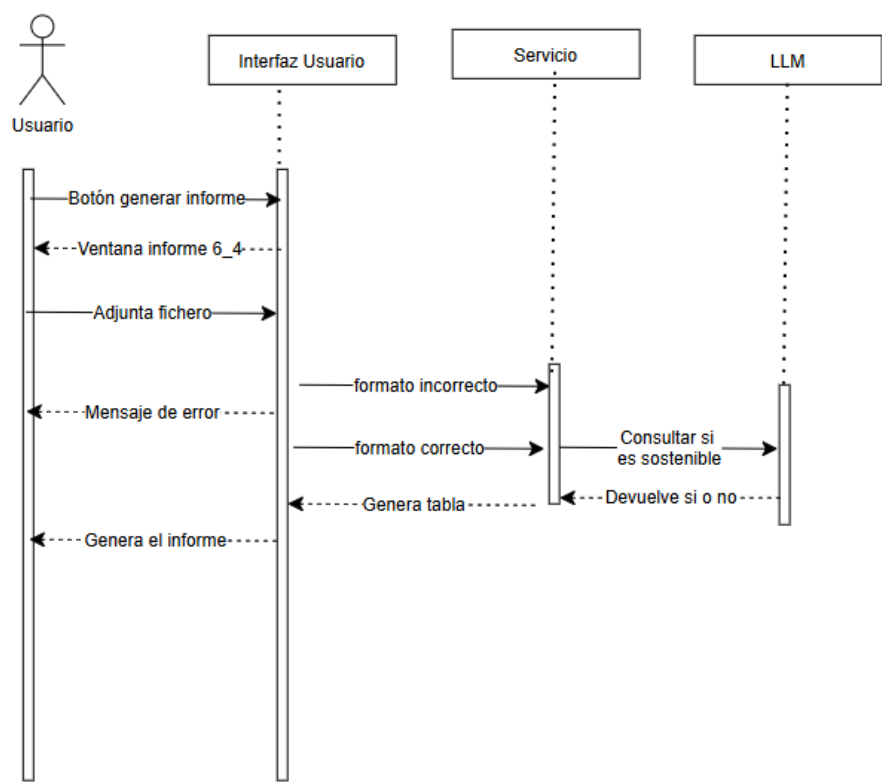


Figura C.9: Diagrama de secuencia de la evidencia 6_4

Number of Scholarly Publications on Sustainability (6_7)

Aquí podemos observar los pasos al ejecutar la evidencia 6_7.

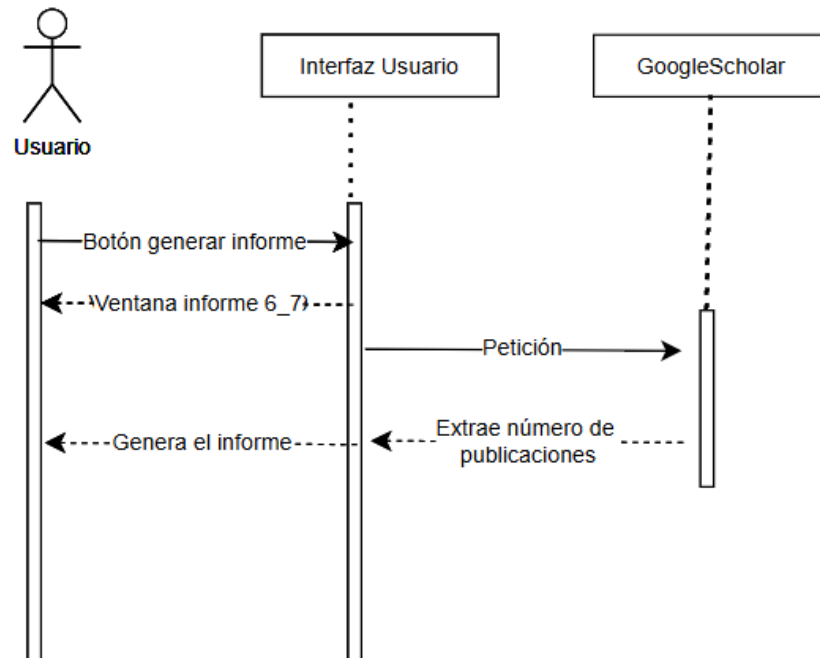


Figura C.10: Diagrama de secuencia de la evidencia 6_7

Number of Events on Environment and Sustainability(6_8)

Aquí podemos observar los pasos al ejecutar la evidencia 6_8.

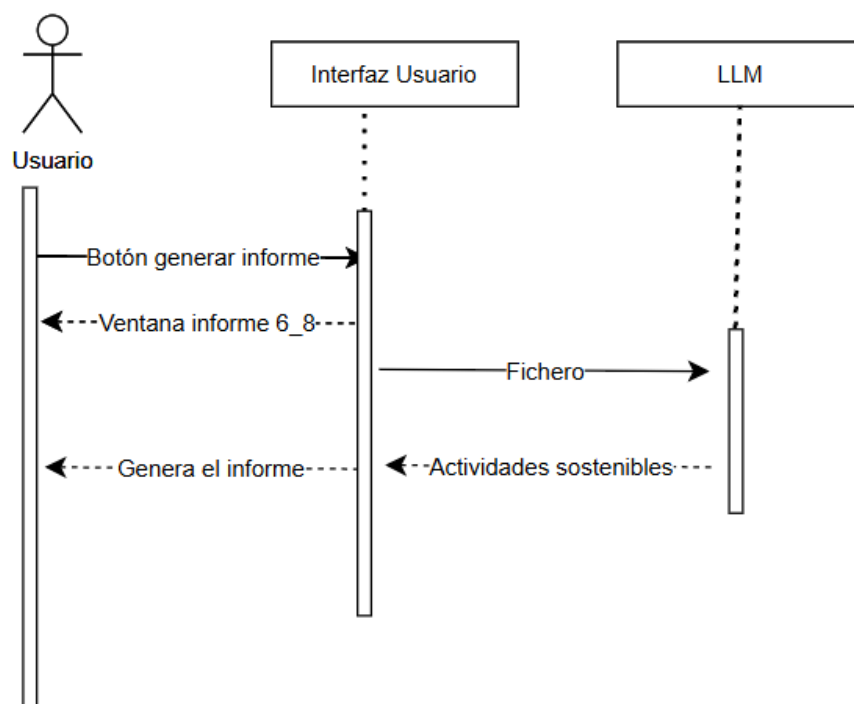


Figura C.11: Diagrama de secuencia de la evidencia 6_8

Apéndice D

Documentación técnica de programación

D.1. Introducción

En este anexo se detalla la estructura del repositorio, el manual del programador y los requisitos necesarios para la instalación y ejecución del proyecto.

D.2. Estructura de directorios

El repositorio del proyecto está organizado de la siguiente manera:

- **/doc/** : Contiene la documentación general del proyecto.
- **/images** : Incluye todas las imágenes empleadas en la documentación.
 - **/images/Castellano/** : Imágenes utilizadas en el manual de usuario en español.
 - **/images/English/** : Imágenes utilizadas en la versión en inglés del manual.
 - **/images/logo/** : Contiene el logotipo utilizado en el archivo README.
- **/src/** : Contiene el código fuente del proyecto, así como los informes generados.

- **/src/generar_informe** : Incluye los scripts de Python responsables de la generación de informes.
 - **/src/generar_informe/UBU_Verde_informes/**: Almacena los informes proporcionados por UBU Verde, utilizados para el informe *Number of Events on Environment and Sustainability (6_8)*.
- **/src/generated_reports/**: Carpeta en la que se almacenan los informes generados automáticamente, siguiendo la estructura mostrada en la figura E.1.
- **/src/instance/**: Contiene las bases de datos locales para usuarios y guías docentes.
- **/src/sostenibilidad/**: Código relacionado con la descarga y análisis de la sostenibilidad de las guías docentes.
 - **/src/sostenibilidad/data/**: Carpeta para almacenar los archivos Excel generados y las guías descargadas.
 - **/src/sostenibilidad/data/guias/**: Contiene las guías docentes descargadas.
- **/src/static/**: Archivos estáticos utilizados en la interfaz (CSS e imágenes).
 - **/src/static/css/**: Hojas de estilo utilizadas en la interfaz.
 - **/src/static/images/**: Imágenes empleadas como fondo o en elementos visuales.
- **/src/templates/**: Archivos HTML que definen la estructura visual de la aplicación.
- **/src/uploads/**: Carpeta generada automáticamente al subir un archivo Excel para el informe *Total Research Funds Dedicated to Sustainability Research (in US Dollars) (6_4)*. Si no se ha subido ningún archivo, esta carpeta no aparecerá.

D.3. Manual del programador

Para utilizar o modificar esta aplicación, se deben tener instalados los siguientes componentes:

Instalación de Python

Python puede descargarse desde el sitio oficial: [2]. Se recomienda instalar la última versión estable.

Instalación de dependencias con pip

En condiciones normales, las dependencias necesarias se instalarán automáticamente desde el archivo `requirements.txt`. Para ello, abra una terminal, acceda al directorio del proyecto, y desde la carpeta `/src`, ejecute:

```
pip install -r requirements.txt
```

Instalación manual de Flask (en caso de error)

Si se produce algún error al instalar desde `requirements.txt`, Flask puede instalarse manualmente con:

```
pip install Flask
```

Esto instalará también las siguientes dependencias asociadas:

- Werkzeug
- Jinja
- MarkupSafe
- ItsDangerous
- Click
- Blinker

Uso de python-dotenv

La aplicación utiliza la librería `python-dotenv` para cargar variables de entorno desde un archivo `.env`. Esta dependencia se instala automáticamente si está incluida en el archivo `requirements.txt`.

En caso de que ocurra algún error durante la instalación, se puede instalar manualmente con el siguiente comando:

```
pip install python-dotenv
```

Instalación de Git

Para clonar y trabajar con el repositorio se necesita Git. Puede descargarse desde: [\[4\]](#).

D.4. Compilación, instalación y ejecución del proyecto

En este apartado se explica cómo obtener el código fuente, instalar sus dependencias y ejecutar el proyecto.

Descargando desde el repositorio

Obtención código fuente

El código fuente se encuentra en un repositorio de GitHub. Para obtener una copia local, siga los siguientes pasos:

1. Cree un directorio donde desee guardar el proyecto.
2. Abra la terminal (Git Bash o similar) en ese directorio.
3. Ejecute el siguiente comando para clonar el repositorio:

```
git clone https://github.com/Lorenah2022/  
GreenMetrics.git
```

Ejecución del proyecto

Acceda desde la terminal a la carpeta `GreenMetrics/src`:

```
cd GreenMetrics/src
```

Instale las dependencias:

```
pip install -r requirements.txt
```

Y ejecute la aplicación:

```
python app.py
```

Abra un navegador y acceda a: <http://127.0.0.1:5000>

Variables de entorno necesarias

Algunas credenciales se almacenan en el archivo `.env` (no incluido en el repositorio por razones de seguridad). Para que la aplicación funcione correctamente, se deben definir las siguientes variables:

- PYTHONUNBUFFERED=1
- DATABASE_URL=sqlite:///database.db
- DATABASE_BINDS=sqlite:///busqueda.db
- GOOGLE_CLIENT_ID=your-google-client-id
- GOOGLE_CLIENT_SECRET=your-google-client-secret
- OAUTHLIB_INSECURE_TRANSPORT=1

Para obtener las credenciales de Google, siga los pasos descritos en: [\[1\]](#) la guía oficial de Google.

Uso con máquina virtual

Si se utiliza la máquina virtual proporcionada, el PIN de acceso es: soyyo13.

D.5. Pruebas del sistema

Para verificar el funcionamiento del sistema, se realizaron pruebas manuales a través de la interfaz de usuario. No se utilizaron marcos de testeo automatizados.

Apéndice E

Documentación de usuario

E.1. Introducción

En esta sección se detallan los requerimientos de la aplicación e indicaciones de como usarla.

E.2. Requisitos de usuarios

De momento la aplicación no se encuentra desplegada por lo que para usarla es necesario usar la máquina virtual que se ha facilitado, donde ya vienen instalados todos los requisitos. Algunos de los requisitos que ya vienen instalados son:

- El framework Flask junto con sus extensiones, utilizadas para el desarrollo de la aplicación web.
- Librerías para la gestión de bases de datos como SQLAlchemy y psycopg2.
- Herramientas de autenticación mediante Google, como Flask-Dance, google-auth, y oauthlib.
- Librerías de procesamiento de documentos como openpyxl, PyPDF2, python-docx, y docx2pdf.
- Utilidades para scraping y automatización como beautifulsoup4, feedparser, y selenium.

- Otras librerías esenciales como pandas, numpy, matplotlib, python-dotenv y tqdm, que permiten el manejo de datos, generación de gráficos y configuración del entorno.

Es necesario tener acceso a internet para poder usar la aplicación.

E.3. Instalación

No es necesario instalar nada para utilizar GreenMetrics, solamente es necesario utilizar un navegador para poder navegar por las distintas interfaces de la aplicación.

E.4. Manual del usuario

Para el manual del usuario, se ha realizado una wiki en castellano, accesible a través del [enlace](#), y su versión en inglés, en el siguiente [enlace](#).

Los informes generados se almacenan en la carpeta *generated_reports*. En su interior, se crean subcarpetas con el nombre de cada tipo de informe, y dentro de estas se organizan los informes por año. Cada subcarpeta anual contiene las versiones del informe correspondientes en formato Word y PDF.

Esta estructura está diseñada para facilitar la localización y el acceso a los informes generados en años específicos.

A continuación, se presenta un esquema [E.1](#) que ilustra la organización de carpetas utilizada para almacenar los informes generados. Esta estructura permite una localización rápida de los informes por tipo y año, manteniendo versiones en formatos PDF y Word.

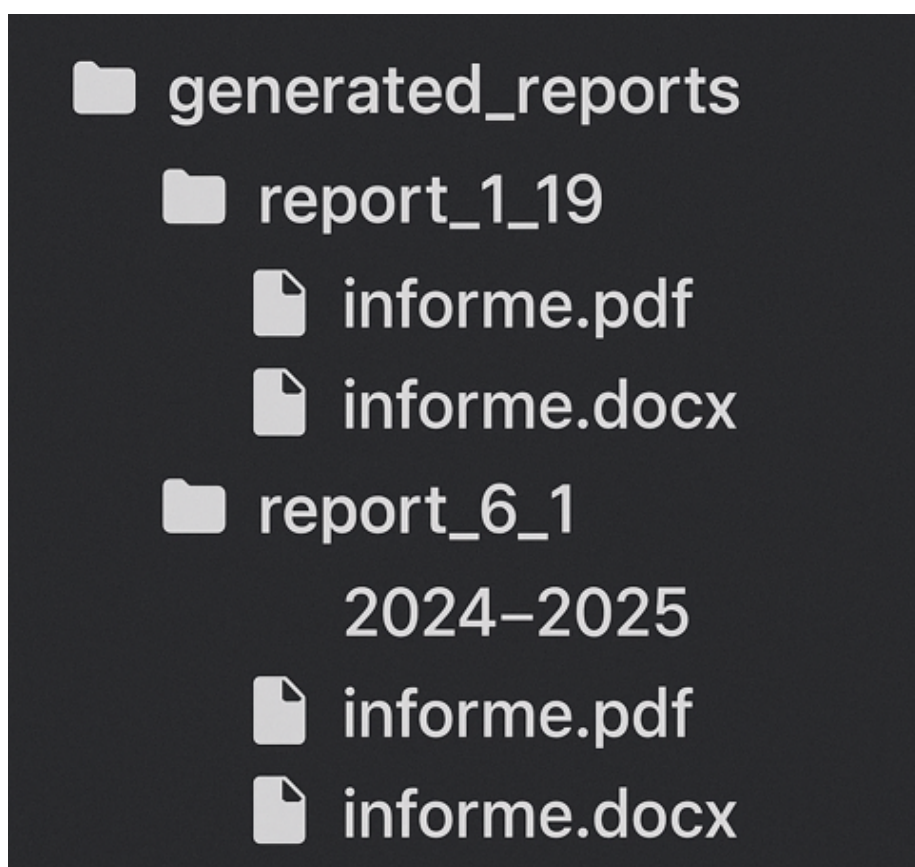


Figura E.1: Esquema carpetas

Apéndice F

Anexo de sostenibilización curricular

F.1. Introducción

Este anexo incluye una reflexión sobre los aspectos de la sostenibilidad abordados en el desarrollo del trabajo de fin de grado. Se presentan las competencias adquiridas durante el desarrollo del sistema de información y su aplicación en el contexto académico y técnico.

Más información sobre sostenibilidad en el ámbito universitario en el documento de la CRUE https://www.crue.org/wp-content/uploads/2020/02/Directrices_Sostenibilidad_Crue2012.pdf.

F.2. Impacto de la sostenibilidad en el proyecto

El sistema de información de apoyo a GreenMetrics ha sido diseñado con el objetivo de mejorar la recopilación y análisis de datos relacionados con la sostenibilidad universitaria , proporcionando métricas clave para evaluar el impacto de iniciativas sostenibles en instituciones educativas.

Durante el desarrollo del proyecto, se han aplicado criterios de eficiencia y responsabilidad ambiental , teniendo en cuenta los siguientes aspectos:

- **Automatización de procesos:** La implementación de métodos automáticos de extracción y procesamiento de guías docentes ha reducido

la necesidad de intervención manual, minimizando el uso de recursos físicos como papel o impresión de documentos.

- **Uso de tecnologías sostenibles:** Se ha utilizado software libre y herramientas optimizadas para reducir el consumo energético asociado al procesamiento de datos.
- **Análisis de competencias en sostenibilidad:** A través de modelos de lenguaje (LLMs), se han identificado asignaturas y programas académicos que incluyen competencias relacionadas con los Objetivos de Desarrollo Sostenible (ODS).
- **Optimización de almacenamiento:** La estructura de bases de datos permite almacenar información de manera eficiente, reduciendo la redundancia y facilitando el acceso a datos históricos sobre sostenibilidad universitaria.

F.3. Competencias adquiridas y aplicadas

El desarrollo del proyecto ha permitido al alumnado profundizar en competencias clave relacionadas con la sostenibilidad y la tecnología:

Competencias técnicas

- **Desarrollo y automatización de procesos:** Implementación de sistemas de extracción y clasificación de datos a partir de documentos académicos en formato PDF.
- **Integración de inteligencia artificial:** Uso de modelos de lenguaje (LLMs) para la identificación de competencias en sostenibilidad.
- **Gestión eficiente de bases de datos:** Uso de Flask-SQLAlchemy para la persistencia y optimización del almacenamiento de información sobre titulaciones y asignaturas.

Competencias en sostenibilidad

- **Análisis de impacto académico:** Evaluación de la presencia de asignaturas orientadas a sostenibilidad dentro de planes de estudio universitarios.

- **Implementación de métricas de sostenibilidad:** Generación de informes sobre programas académicos que incorporan competencias relacionadas con los ODS.
- **Optimización de recursos digitales:** Uso de herramientas digitales para reducir el consumo de materiales impresos en la evaluación de iniciativas sostenibles.

F.4. Conclusión

El desarrollo de este trabajo de fin de grado ha permitido no solo la aplicación de conocimientos técnicos en inteligencia artificial y gestión de datos, sino también la integración de criterios de sostenibilidad universitaria en el proceso. Gracias a la automatización de la recopilación y análisis de datos, el sistema ha sido optimizado para ofrecer métricas clave sobre la implementación de competencias sostenibles en programas de estudio universitarios.

Las competencias adquiridas a lo largo del trabajo han reforzado la importancia de la innovación tecnológica aplicada a la sostenibilidad, contribuyendo al desarrollo de metodologías que faciliten la transición hacia universidades más comprometidas con el medio ambiente y la responsabilidad social.

Este anexo se alinea con los principios establecidos en el documento de la CRUE sobre sostenibilidad académica, reforzando la importancia de la integración de competencias sostenibles en el ámbito universitario.

Bibliografía

- [1] Google Developers. Crear credenciales OAuth 2.0 — Google Workspace for Developers. <https://developers.google.com/workspace/guides/create-credentials?hl=es-419#oauth-client-id>, 2025. [En línea; consultado el mayo-2025].
- [2] Python Software Foundation. Descargas de Python - Sitio oficial. <https://www.python.org/downloads/>, 2025. [En línea; consultado el mayo-2025].
- [3] Seguridad Social. Cotización y recaudación de los trabajadores - seguridad social. <https://www.seg-social.es/wps/portal/wss/internet/Trabajadores/CotizacionRecaudacionTrabajadores/36537#36538/>, 2024. [Internet; consultado el mayo- 2025].
- [4] Software Freedom Conservancy. Git - Distributed Version Control System. <https://git-scm.com/>, 2025. [En línea; consultado el mayo-2025].
- [5] Wikipedia. Modelo vista controlador. <https://es.wikipedia.org/w/index.php?title=Modelo> [En línea; consultado el marzo-2025].
- [6] Wikipedia. Facade (patrón de diseño). <https://es.wikipedia.org/wiki/Facade> [En línea; consultado el junio-2025].