



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Gº en Ingeniería en Informática



TFG Ingeniería Informática:
Ububooknet



Presentado por Luis Miguel Cabrejas
en Burgos el 1 julio de 2019
Tutores D. José Manuel Galán Ordax
y D. Luis Rodrigo Izquierdo Millán

Índice General

Índice General	1
Índice de figuras	2
Índice de tablas	4
Apéndice A. Planificación.....	5
A.1. Introducción.....	5
A.2. Planificación temporal.....	6
A.3. Estudio económico	24
A.4. Tratamiento de licencias.....	26
Apéndice B. Especificación de Requisitos.....	27
B.1. Introducción.....	27
B.2. Objetivos generales.....	27
B.3. Catálogo de requisitos.....	28
B.4. Especificación de requisitos	30
Apéndice C. Especificación de Diseño	40
C.1. Introducción.....	40
C.2. Diseño de datos.....	40
C.3. Diseño arquitectónico	41
C.4. Diseño de interfaces.....	43
Apéndice D. Documentación Técnica de Programación.....	46
D.1. Introducción.....	46
D.2. Estructura de directorios.....	46
D.3. Manual del programador	47
D.4. Pruebas Unitarias.....	51
D.5. Scraping para obtener diccionarios de personajes.....	51
Apéndice E. Documentación de Usuario	53
E.1. Introducción	53
E.2. Requisitos de usuarios.....	53
E.3. Instalación	53
E.4. Manual de Usuario.....	55

Índice de figuras

Figura A 1 Gráfico burndown Sprint 1	6
Figura A 2 Gráfico burndown Sprint 2	8
Figura A 3 Gráfico burndown Sprint 3	9
Figura A 4 Gráfico burndown Sprint 4	10
Figura A 5 Gráfico burndown Sprint 5	12
Figura A 6 Gráfico burndown Sprint 6	13
Figura A 7 Gráfico burndown Sprint 7	14
Figura A 8 Gráfico burndown Sprint 8	15
Figura A 9 Gráfico burndown Sprint 9	17
Figura A 10 Gráfico burndown Sprint 10	18
Figura A 11 Gráfico burndown Sprint 11	19
Figura A 12 Gráfico burndown Sprint 12	21
Figura A 13 Gráfico burndown Sprint 13	22
Figura A 14 Gráfico burndown Sprint 14	23
Figura C 1 Diagrama de paquetes	42
Figura C 2 Diagrama de clases	42
Figura C 3 Prototipo: Inicio 1	43
Figura C 4 Prototipo: Menú 2	43
Figura C 5 Prototipo: Diccionarios Automáticos 2.A	44
Figura C 6 Prototipo: Ver Diccionario 2.B	44
Figura C 7 Prototipo: Modificar Diccionario 2.C	45
Figura D 1 Ejecución de la aplicación	48
Figura D 2 Crear nueva app heroku	49
Figura D 3 Información para crear la app	49
Figura D 4 Vincular con GitHub	49
Figura D 5 Selección de la rama	50
Figura D 6 Diccionario idiomas	51
Figura D 7 Código a comentar y descomentar	51
Figura D 8 Ejemplo Wikipedia Juego de Tronos	52
Figura D 9 Ejemplo Wikipedia Harry Potter	52
Figura D 10 Ejemplo Fandom Juego de Tronos	52
Figura D 11 Ejemplo Fandom Harry Potter	53
Figura E 1 Botones del navegador	55
Figura E 2 Menú de la aplicación	55
Figura E 3 Interfaz de inicio	56
Figura E 4 Interfaz de diccionarios automáticos	57
Figura E 5 Interfaz de importar diccionario	57

Figura E 6 Ejemplo de archivo csv de personajes	58
Figura E 7 Interfaz de obtener diccionario	58
Figura E 8 Vista de la página donde realizar scraping	59
Figura E 9 Interfaz de modificar diccionario	60
Figura E 10 Interfaz de añadir personaje.....	61
Figura E 11 Interfaz de eliminar personaje	61
Figura E 12 Interfaz de juntar personajes.....	62
Figura E 13 Interfaz añadir referencia.....	63
Figura E 14 Interfaz de eliminar referencia.....	63
Figura E 15 Interfaz modificar id personaje.....	64
Figura E 16 Interfaz parámetros adicionales	65
Figura E 17 Interfaz visualizar red	65
Figura E 18 Interfaz generación de informes	67

Índice de tablas

Tabla A. 1 Story points del proyecto.....	5
Tabla A. 2 Etiquetas Sprint 1	7
Tabla A. 3 Etiquetas Sprint 2	8
Tabla A. 4 Etiquetas Sprint 3	9
Tabla A. 5 Etiquetas Sprint 4	11
Tabla A. 6 Etiquetas Sprint 5	12
Tabla A. 7 Etiquetas Sprint 6	13
Tabla A. 8 Etiquetas Sprint 7	15
Tabla A. 9 Etiquetas Sprint 8	16
Tabla A. 10 Etiquetas Sprint 9	17
Tabla A. 11 Etiquetas Sprint 10	19
Tabla A. 12 Etiquetas Sprint 11	20
Tabla A. 13 Etiquetas Sprint 12	21
Tabla A. 14 Etiquetas Sprint 13	22
Tabla A. 15 Etiquetas Sprint 13	23
Tabla A. 16 Licencias de dependencias	26
Tabla B. 1 Introducir ePub	30
Tabla B. 2 Crear diccionario personajes	31
Tabla B. 3 Importar diccionario	31
Tabla B. 4 Obtener diccionario	32
Tabla B. 5 Vaciar diccionario.....	32
Tabla B. 6 Añadir personaje.....	33
Tabla B. 7 Eliminar personaje.....	33
Tabla B. 8 Juntar personajes.....	34
Tabla B. 9 Añadir nueva referencia.....	34
Tabla B. 10 Eliminar referencias de personajes	35
Tabla B. 11 Modificar id diccionario	35
Tabla B. 12 Exportar diccionario	36
Tabla B. 13 Obtener posiciones personajes.....	36
Tabla B. 14 Generar red	37
Tabla B. 15 Exportar red	37
Tabla B. 16 Visualizar red.....	38
Tabla B. 17 Generar informes	38
Tabla B. 18 Internacionalización.....	39

Apéndice A. Planificación

A.1. Introducción

En esta sección se detalla la planificación utilizada para realizar el proyecto, indicando las tareas que se realizan, junto con el sprint en el que se realizan y la duración tanto del sprint como la estimación de duración de cada tarea.

Para gestionar las estimaciones temporales de las tareas, los sprints y el tipo de tarea que se está realizando utilizamos la herramienta ZenHub.

Lo primero que realizamos con ZenHub es añadir algunos tipos de tareas que por defecto no vienen añadidos, se añaden: configuration, design, documentation, feature y research.

A la hora de estimar el tiempo que se va a tardar en realizar una tarea ZenHub ofrece distintos valores conocidos como story points. Se asigna un valor temporal a cada uno de estos valores y se obtiene este resultado:

Story Pons	Estimación temporal
1	45 minutos
2	90 minutos
3	3 horas
5	5 horas
8	8 horas
13	16 horas
21	24 horas
40	40 horas

Tabla A. 1 Story points del proyecto

A.2. Planificación temporal

A continuación, se detallan todos los sprints realizados a lo largo del proyecto estableciendo sus objetivos, las tareas desarrolladas, el tipo al que pertenecen y el grado de cumplimiento y distribución temporal mediante los gráficos de velocidad. Se incluye en cada sprint el enlace en el repositorio en ZenHub para facilitar la consulta.

Sprint 1.

Se inicia el proyecto.

Se lee la guía para la realización del TFG y se crea y configura el repositorio. Después de esto se investiga sobre las herramientas a utilizar durante el desarrollo del proyecto y sobre la estructura de un ePub. Finalmente se documentan estas decisiones.

Link : [Sprint 1](#)

Fecha inicio / fin : 14 ene 2019 / 25 ene 2019

Objetivo : Inicio del proyecto. Configuración del repositorio y tareas de formación.

Etiquetas : Configuration, Research, Documentation.

Story Points : 20

Gráfico burndown.

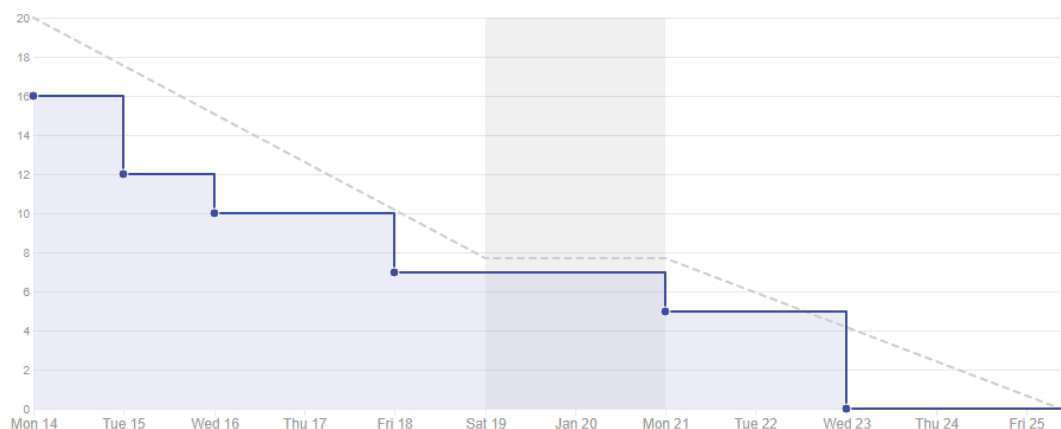


Figura A 1 Gráfico burndown Sprint 1

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#1	Crear y configurar repositorio							
#2	Elegir editor de texto para la memoria							
#3	Elegir gestor de referencias							
#4	Elegir herramientas de prototipado							
#5	Búsqueda de estudios relacionados							
#6	Investigar sobre la estructura de un ePub							
#7	Elegir IDE							
#8	Añadir archivos memoria							
#9	Buscar herramientas para parsear ePubs							
#10	Documentar la investigación realizada en el primer sprint							

Tabla A. 2 Etiquetas Sprint 1

Sprint 2.

Se sigue investigando acerca de herramientas a utilizar en el proyecto y se implementan funciones para trabajar con el diccionario de personajes.

Link : [Sprint 2](#)

Fecha inicio / fin : 25 ene 2019 / 31 ene 2019

Objetivo : Realizar tareas de investigación e implementar las primeras funcionalidades del proyecto.

Etiquetas : Configuration, Research, Documentation, Test, Bug, Feature.

Story Points : 26

Gráfico burndown.

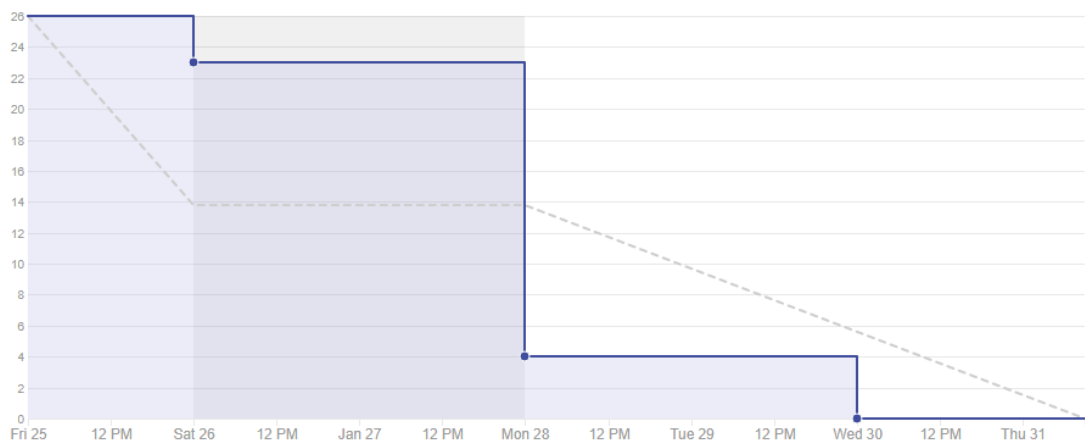


Figura A 2 Gráfico burndown Sprint 2

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#11	Buscar herramientas para parsear texto en Python							
#12	Configurar IDE							
#13	Crear pruebas unitarias							
#14	Encontrar nombres de personajes							
#15	Generar histograma de personajes							
#16	Juntar personajes							
#17	Eliminar personajes							
#18	Añadir personaje							
#19	Estudiar interactividad en Python							
#20	Documentar la investigación realizada en el primer sprint							
#21	Error en las pruebas unitarias creadas							
#22	Ampliar pruebas unitarias para nuevas funciones							
#23	Añadir nueva forma de referirse a un personaje							
#24	Eliminar forma de referirse a un personaje							

Tabla A. 3 Etiquetas Sprint 2

Sprint 3.

Se investiga sobre interfaces para la aplicación y se añade la lectura de ficheros ePub.

Link : [Sprint 3](#)

Fecha inicio / fin : 31 ene 2019 / 8 feb 2019

Objetivo : Investigación sobre interfaces e implementar nuevas funcionalidades

Etiquetas : Research, Feature, Test.

Story Points : 12

Gráfico burndown.

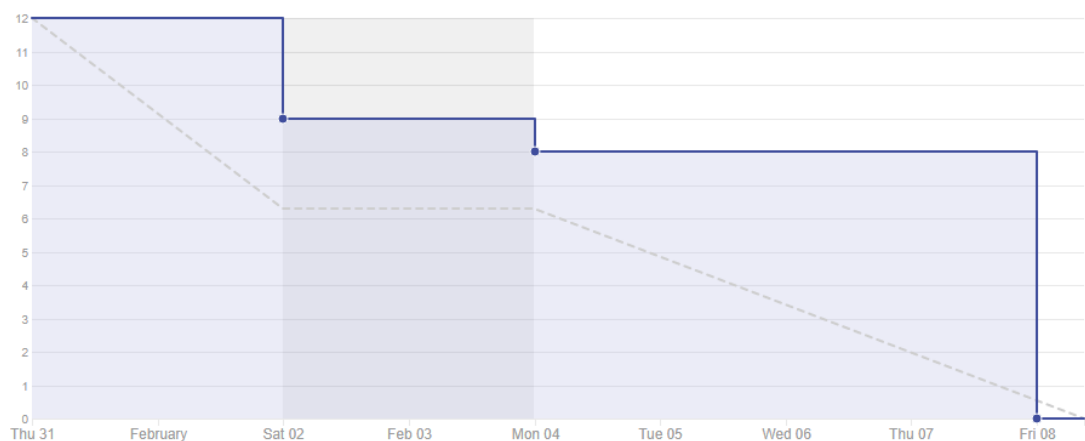


Figura A 3 Gráfico burndown Sprint 3

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#25	Implementar lectura de ePub							
#26	Estudiar distintas herramientas de interfaces							
#27	Añadir test para nuevas funcionalidades							

Tabla A. 4 Etiquetas Sprint 3

Sprint 4.

Se sigue investigando sobre la interfaz y se desarrollan más funcionalidades del proyecto, enfocándose en la parte de generación de la red a partir de un diccionario.

Link : [Sprint 4](#)

Fecha inicio / fin : 8 feb 2019 / 25 feb 2019

Objetivo : Investigación sobre la interfaz y desarrollo de más funcionalidades del proyecto

Etiquetas : Configuration, Research, Documentation, Test, Feature.

Story Points : 21

Gráfico burndown.

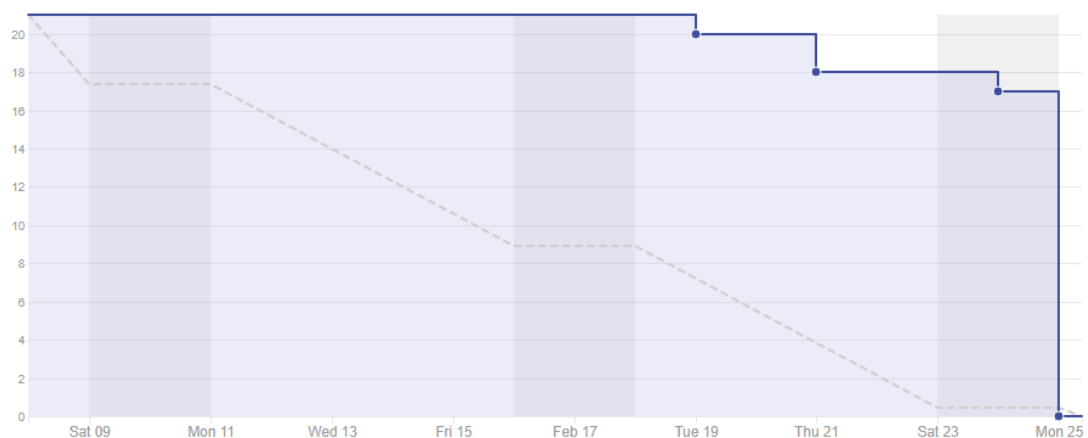


Figura A 4 Gráfico burndown Sprint 4

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#28	Implementar la función de parsear el texto dado un diccionario							
#29	Generar una matriz de adyacencia con los datos obtenidos de parsear con diccionario							

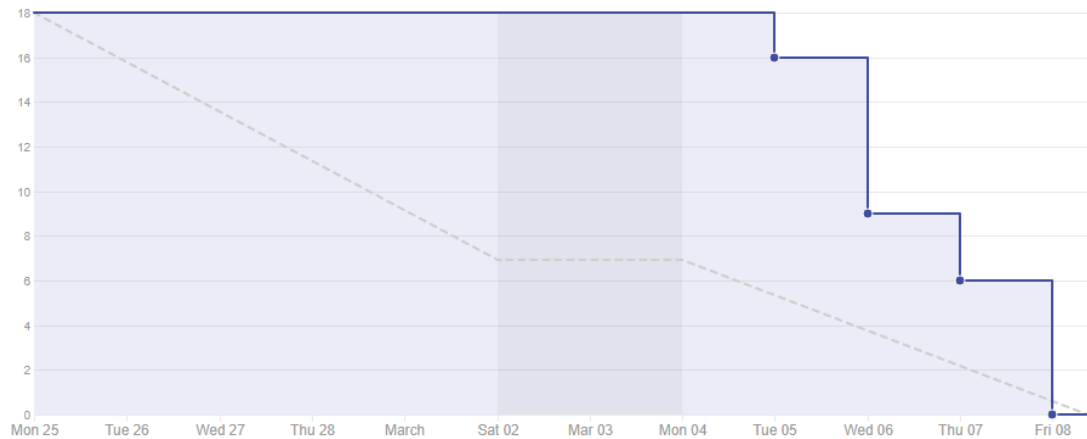


Figura A 5 Gráfico burndown Sprint 5

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#37	Implementar un método para importar un diccionario de personajes							
#38	Implementar un método para exportar un diccionario de personajes							
#39	Investigar el web scraping							
#40	Implementar un método para obtener un diccionario mediante web scraping							
#41	Reorganizar el código para solucionar posibles bugs							
#42	Añadir comentarios al código generado hasta ahora							
#43	Investigar más opciones para la interfaz							

Tabla A. 6 Etiquetas Sprint 5

Sprint 6.

Se realiza un prototipo de interfaz y se implementa la interfaz, además de corregir un bug de los lexer utilizados en crear un diccionario de personajes automáticamente y en obtener las posiciones de los personajes en el texto.

Link : [Sprint 6](#)

Fecha inicio / fin : 8 mar 2019 / 15 mar 2019

Objetivo : Implementar la interfaz gráfica.

Etiquetas : Documentation, Bug, Feature.

Story Points : 14

Gráfico burndown.

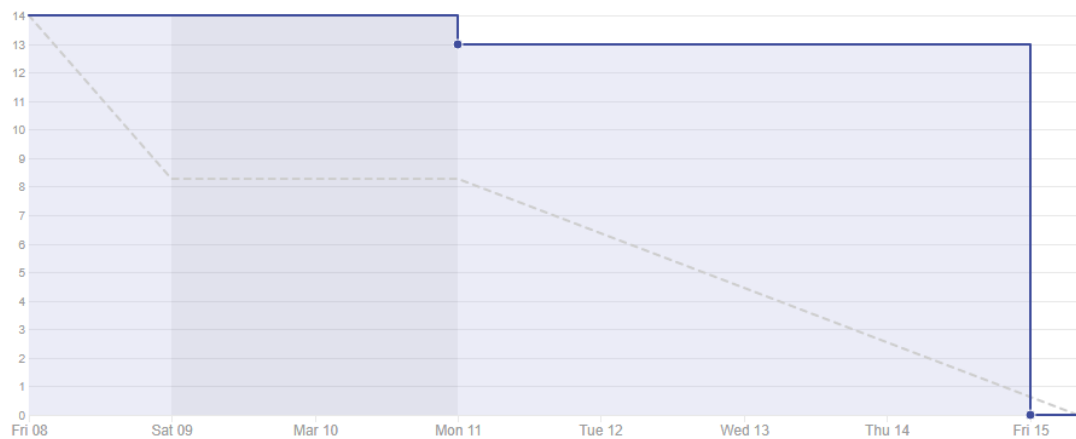


Figura A 6 Gráfico burndown Sprint 6

Etiquetas.

		bug	configuration	design	documentation	feature	research	test
Issues	Titulo							
#44	Bug en los lexer en la creación automática de diccionario de personajes y detección de posición							
#45	Diseñar la interfaz							
#46	Implementar la interfaz gráfica							
#47	Implementar la selección automática de etiqueta							
#48	Documentar las decisiones de diseño realizadas hasta ahora							

Tabla A. 7 Etiquetas Sprint 6

Sprint 7.

Se corrigen distintos bugs y se realizan modificaciones en los métodos y la interfaz para lanzar la primera versión de la aplicación.

Link : [Sprint 7](#)
Fecha inicio / fin : 15 mar 2019 / 29 mar 2019
Objetivo : Publicar la primera versión del proyecto.
Etiquetas : Bug, Feature.
Story Points : 18

Gráfico burndown.

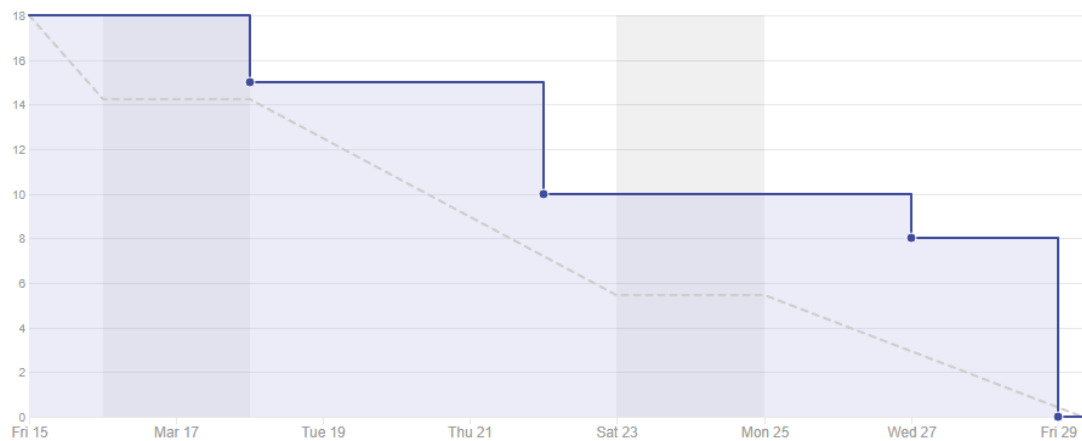


Figura A 7 Gráfico burndown Sprint 7

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#49	Reorganizar la interfaz gráfica para hacerla más sencilla de utilizar							
#50	Solucionar bug en la función para obtener la posición de los personajes							
#51	Bug en la creación automática de personajes							
#52	Modificar los métodos para poder informar cuando los datos que introduzca no sean correctos							

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#57	Permitir al usuario modificar los identificadores de los personajes							
#58	Modificar la creación y la obtención del diccionario de personajes para que no pongan ids numéricas							
#59	Bug al leer el texto de un archivo ePub y al crear el diccionario							
#60	Permitir borrar varios personajes a la vez a través de la interfaz							
#61	Añadir botón para que el usuario pueda exportar su diccionario de personajes							
#62	Conservar la configuración de d3 al cambiar de ventana							
#63	Añadir barra de progresión a procesos que tardan en ejecutarse							
#64	Ordenar los personajes por el número de apariciones							
#65	Implementar método para juntar listas manteniendo el orden							

Tabla A. 9 Etiquetas Sprint 8

Sprint 9.

Se mejora la representación gráfica de la aplicación y se añaden nuevas funcionalidades como el drag & drop u ordenar los personajes por el número de apariciones o alfabéticamente.

Link	: Sprint 9
Fecha inicio / fin	: 10 abr 2019 / 24 abr 2019
Objetivo	: Mejorar usabilidad y añadir funcionalidades
Etiquetas	: Bug, Feature, Design.
Story Points	: 23

Gráfico burndown.

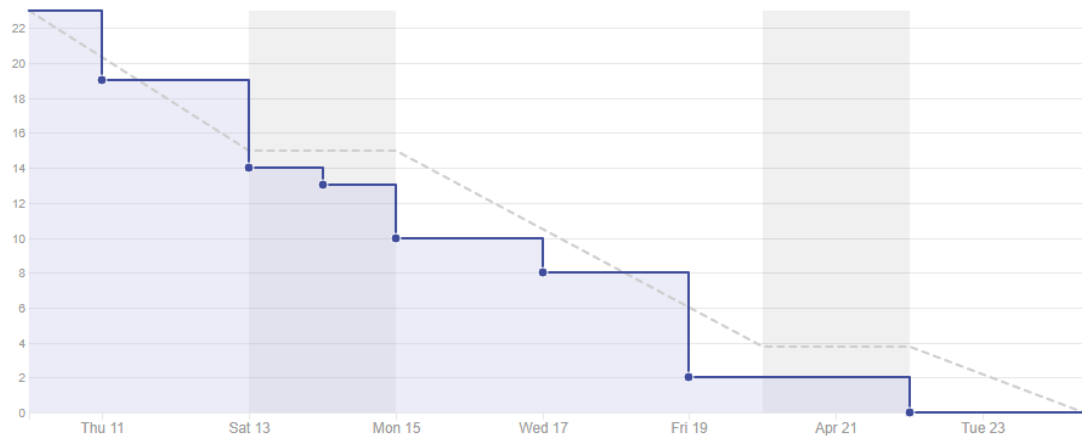


Figura A 9 Gráfico burndown Sprint 9

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#66	Permitir juntar varios personajes a la vez a través de la interfaz							
#67	Permitir juntar personajes mediante drag&drop							
#68	Permitir eliminar varias referencias de distintos personajes a la vez a través de la interfaz							
#69	Permitir al usuario ordenar el diccionario de personajes por orden alfabético o número de apariciones							
#70	Permitir al usuario tener en cuenta los capítulos a la hora de hallar las relaciones entre personajes							
#71	Seguir integrando css en la interfaz							
#72	Añadir identificador de workflow							
#73	Permitir al usuario seguir modificando el diccionario de personajes después de hallar las posiciones de los personajes							
#74	Bug al generar el grafo							
#75	Refactorizar código							

Tabla A. 10 Etiquetas Sprint 9

Sprint 10.

Se internacionaliza la página web mediante Flask-Babel y se añaden nuevas funcionalidades, como exportar la red generada, además de mejorar el css de la aplicación

Link : [Sprint 10](#)

Fecha inicio / fin : 30 abr 2019 / 8 may 2019

Objetivo : Añadir internacionalización y mejorar el diseño de la interfaz

Etiquetas : Research, Feature, Design.

Story Points : 17

Gráfico burndown.

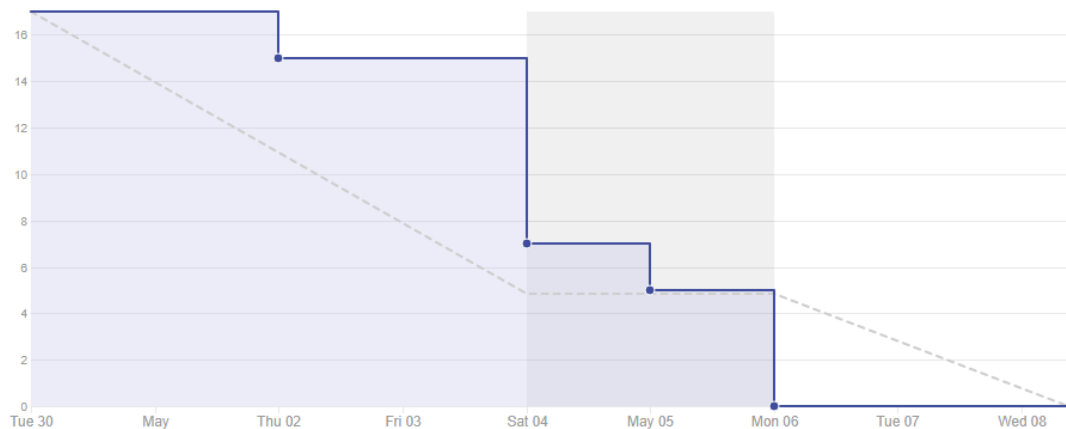


Figura A 10 Gráfico burndown Sprint 10

Etiquetas.

Issues	Titulo			bug	configuration	design	documentation	feature	research	test
#76	Investigar	sobre	cómo							
	internacionalizar	páginas web								

#77	Permitir la visualización de la página tanto en inglés como en español	
#78	Modificar el código para que las comprobaciones de datos las haga el cliente y no el servidor	
#79	Permitir exportar la red generada con los parámetros pasados	
#80	Seguir mejorando la apariencia de la aplicación	

Tabla A. 11 Etiquetas Sprint 10

Sprint 11.

Se investiga en como permitir la concurrencia en Flask y se implementa la concurrencia en la aplicación.

Link : [Sprint 11](#)
Fecha inicio / fin : 30 abr 2019 / 8 may 2019
Objetivo : Permitir concurrencia
Etiquetas : Bug, Research, Feature.
Story Points : 16

Gráfico burndown.

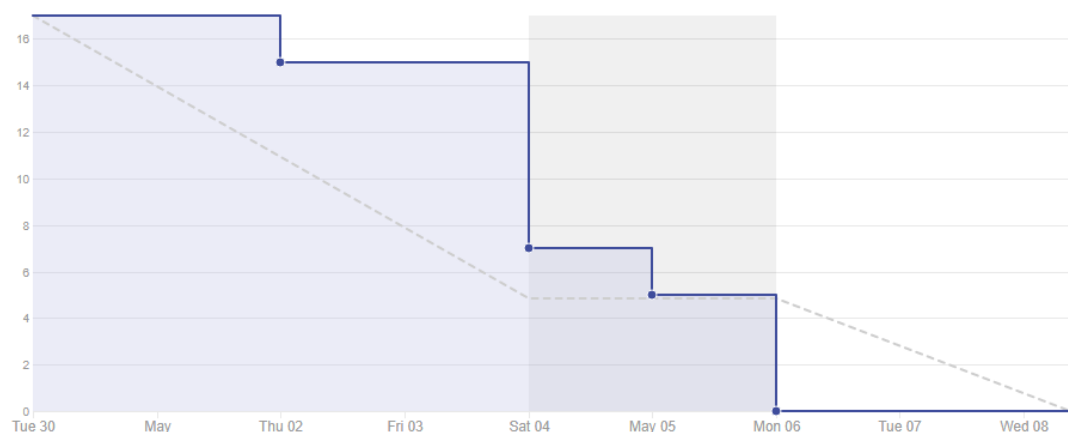


Figura A 11 Gráfico burndown Sprint 11

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#81	Realizar las ordenaciones en el lado del cliente							
#82	Investigar la concurrencia en Flask							
#83	Implementar concurrencia en la aplicación							
#84	Bug por que se le permite introducir textos vacíos como válidos en la aplicación							

Tabla A. 12 Etiquetas Sprint 11

Sprint 12.

Se redacta la memoria y se permite generar informes en la aplicación.

Link : [Sprint 12](#)
Fecha inicio / fin : 6 jun 2019 / 14 jun 2019
Objetivo : Redactar la memoria
Etiquetas : Documentation, Feature
Story Points : 34

Gráfico burndown.

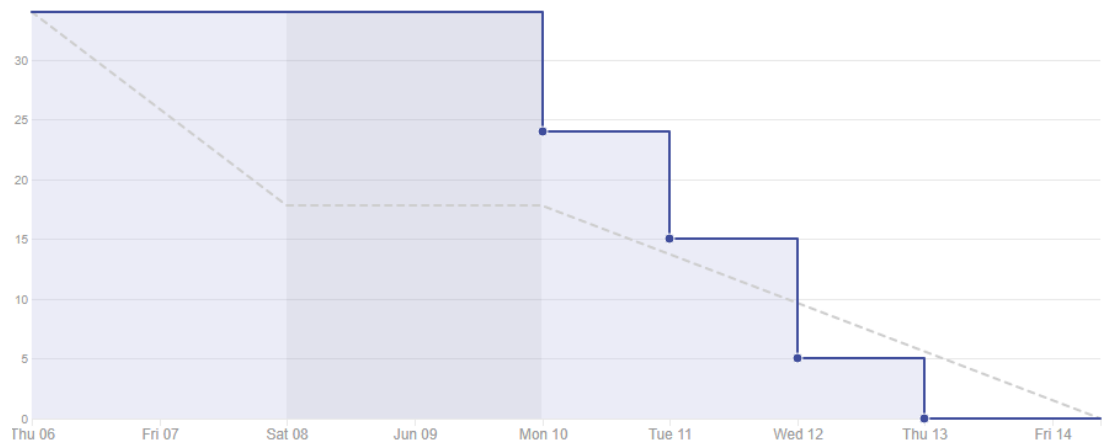


Figura A 12 Gráfico burndown Sprint 12

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#85	Escribir sobre los conceptos teóricos							
#86	Redactar los aspectos relevantes							
#87	Terminar de redactar el apartado de Técnicas y Herramientas							
#88	Actualizar documentación sobre trabajos relacionados							
#89	Redactar la introducción del proyecto							
#90	Redactar los objetivos del proyecto							
#91	Redactar las conclusiones del proyecto							
#92	Añadir la funcionalidad de generar informes							

Tabla A. 13 Etiquetas Sprint 12

Sprint 13.

Se despliega la aplicación en Heroku y se comienzan a redactar los anexos

Link : [Sprint 13](#)

Fecha inicio / fin : 14 jun 2019 / 19 jun 2019

Objetivo : Desplegar la aplicación

Etiquetas : Documentation, Feature

Story Points : 23

Gráfico burndown.

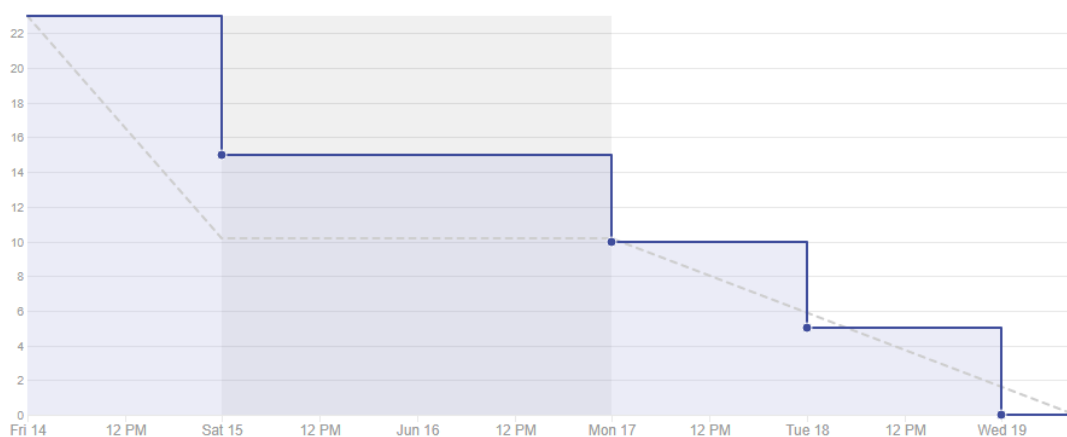


Figura A 13 Gráfico burndown Sprint 13

Etiquetas.

		bug	configuration	design	documentation	feature	research	test
Issues	Titulo							
#93	Desplegar aplicación							
#94	Crear la guía de usuario							
#95	Crear la guía de programador							
#98	Redactar la especificación de diseños							

Tabla A. 14 Etiquetas Sprint 13

Sprint 14.

Se finalizan los anexos y se corrigen fallos en la documentación. También se mejora la generación de informe añadiendo imágenes.

Link : [Sprint 14](#)

Fecha inicio / fin : 19 jun 2019 / 28 jun 2019

Objetivo : Finalizar la documentación y mejorar la generación de informes

Etiquetas : Documentation, Feature, Design, Test

Story Points : 21

Gráfico burndown.

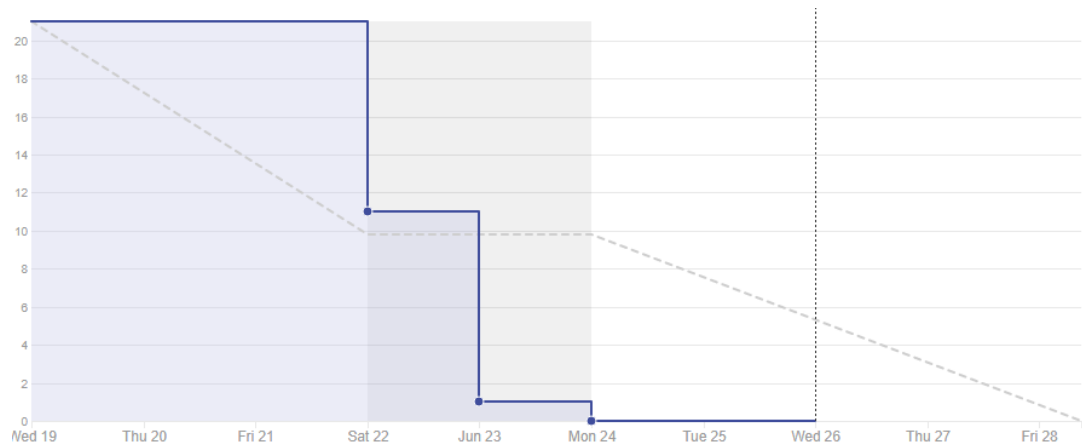


Figura A 14 Gráfico burndown Sprint 14

Etiquetas.

Issues	Titulo	bug	configuration	design	documentation	feature	research	test
#96	Redactar la planificación							
#97	Redactar la especificación de requisitos							
#99	Mejorar la generación de informes							
#100	Actualizar test unitarios							
#101	Crear pestaña de Acerca De							

Tabla A. 15 Etiquetas Sprint 13

A.3. Estudio económico

En este apartado se detallan los costes que derivados de llevar a cabo este proyecto.

Viabilidad económica.

A pesar de ser un proyecto con uno de los objetivos principales docentes, de este proyecto se podrían derivar ingresos ofreciendo servicios de suscripción a editoriales de libros para que puedan usar el software con distintos libros antes de ser publicados, y gracias a la información que se puede extraer de la red generada, estimar si el libro tendrá éxito o no y a qué público deben enfocar su venta aparte de su uso generalizado como generador de características interpretables en filtros colaborativos y otros tipos de sistemas de recomendación.

Costes software.

Durante el desarrollo del proyecto no se utiliza ningún software con licencia de pago por lo que los costes software suponen un total de 0€.

Costes hardware.

A la hora de desplegar la aplicación se utiliza un servicio de hosting como es Heroku, al no ser una aplicación altamente utilizada el servicio nos es gratuito, en caso de que la aplicación fuese altamente concurrida o necesitásemos mayor capacidad de cómputo Heroku ofrece unos servicios que van desde los 25 a los 500 euros al mes,

Costes de recursos humanos.

Durante el transcurso del proyecto se estima que el alumno tutorado ha trabajado unas 16h por semana, y que se le paga 10€/h. Esto hace un coste mensual en salario bruto de:

$$16h/semana \times 10€/h \times 4semana/mes = 640€/mes$$

Esto sería el salario bruto del alumno tutorado, a lo que hay que añadir los impuestos que ha pagado la empresa previamente, 23.6% de contingencias comunes, 5.5% de desempleo, 0.6% de formación profesional y un 0.2% de FOGASA. Los costes totales para la empresa son:

$$\frac{640€/mes \times 100}{100 - (23.6 + 5.5 + 0.6 + 0.2)} = 843€/mes$$

Además del alumno tutorado, en el desarrollo del proyecto han trabajado 2 cotutores que se estima que han trabajado 1h por semana cada uno, y al tener más conocimientos se les paga 36€/h. Esto supone un coste mensual en salario bruto de:

$$2 \times (1h/semana \times 36€/hora \times 4semana/mes) = 288€/mes$$

Al igual que antes, la empresa también debe pagar impuestos, lo que supone un coste mensual total de:

$$\frac{288€/mes \times 100}{100 - (23.6 + 5.5 + 0.6 + 0.2)} = 411€/mes$$

La suma de todos los salarios supone a la empresa un coste mensual total de:

$$913€/mes + 137€/mes = 1040€/mes$$

Al tener el proyecto una duración de 6 meses, eso hace que los costes totales del desarrollar el proyecto sean:

$$6mes \times 1040€/mes = 6240€$$

A.4. Tratamiento de licencias

Antes de establecer una licencia para el proyecto se debe evaluar las licencias del software que estamos utilizando, y de esta forma establecer una licencia compatible.

El software utilizado en el proyecto y sus licencias son las siguientes:

Módulo	Versión	Licencia
BeautifulSoup4	4.7.1	MIT
Flask	1.0.2	BSD
Flask-Babel	0.12.2	BSD
html5lib	1.0.1	MIT
lxml	4.3.1	BSD
matplotlib	3.0.2	PSF
networkx	2.2	BSD
network_styling_with_d3	-	MIT
ply	3.11	BSD
numpy	1.14.3	BSD
scipy	1.1.0	BSD

Tabla A. 16 Licencias de dependencias

Las licencias del software utilizado en el proyecto permiten su uso bajo total libertad.

Se decide utilizar la licencia MIT como licencia del proyecto, permitiendo su uso, copia, modificación, publicación, distribución, sublicenciar, o venta.

Apéndice B. Especificación de Requisitos

B.1. Introducción

En esta sección se describen los objetivos y funcionalidades que se pretenden desarrollar con la realización del proyecto.

B.2. Objetivos generales

Los objetivos generales que se pretenden conseguir con la realización del proyecto son los siguientes:

- Desarrollar una aplicación web que permita generar una red de interacciones entre los personajes de cualquier ePub.
- Permitir al usuario generar un diccionario de personajes de forma completamente automática mediante:
 - Búsqueda de palabras en mayúsculas en el ePub
 - Importación de un diccionario en csv
 - Scraping de una wiki sobre el libro
- Permitir al usuario modificar su diccionario de personajes de distintas formas:
 - Añadiendo nuevos personajes o añadiendo nuevas referencias a un personaje ya existente
 - Eliminando personajes ya existentes o eliminando referencias de un personaje ya existente
 - Juntar varios personajes en uno solo
 - Modificar la id de los personajes
- Permitir al usuario visualizar la red con distintos parámetros para generarla
- Permitir al usuario obtener una serie de información relevante acerca de la red

B.3. Catálogo de requisitos

Requisitos funcionales

- **RF-1 Introducción de fichero ePub.** Cualquier acceso a secciones de la aplicación que no sea el inicio, sin haber introducido un fichero ePub se redirigirán a la página de inicio.
- **RF-2 Validación de ePub.** Se comprobará que efectivamente el archivo introducido tiene la estructura de un archivo ePub.
- **RF-3 Obtención del texto de ePub.** Una vez validado el archivo ePub introducido, se extraerá el texto que compone el libro y se eliminará el archivo ePub.
- **RF-4 Creación automática de un diccionario de personajes.** Se podrá crear un diccionario de personajes a partir de todas las palabras en mayúsculas del ePub subido.
- **RF-5 Importación de un diccionario de personajes.** Se podrá generar un diccionario de personajes a partir de un archivo csv que siga con la estructura marcada en el manual de usuario.
- **RF-6 Obtención de un diccionario de personajes.** Se podrá generar un diccionario de personajes introduciendo el enlace a una wiki pre-establecida.
- **RF-7 Vaciar el diccionario de personajes.** Se podrá eliminar completamente el diccionario de personajes actual para empezar desde cero.
- **RF-8 Añadir personajes.** Se podrán añadir nuevos personajes mientras que su identificador sea distinto al del resto de personajes
- **RF-9 Eliminar personajes.** Se podrán eliminar personajes pertenecientes al diccionario de personajes
- **RF-10 Juntar personajes.** Se podrán juntar varios personajes para formar un único personaje que agrupe las referencias de ambos personajes.
- **RF-11 Añadir referencias a personajes existentes.** Se podrán añadir nuevas referencias a personajes que ya existan en el diccionario de personajes.
- **RF-12 Eliminar referencias a personajes existentes.** Se podrán eliminar referencias de los distintos personajes que formen el personaje.

- **RF-13 Modificar la id de los personajes.** Se podrá modificar la id de los personajes a otra que no esté ya asignada.
- **RF-14 Exportar diccionario.** Se podrá exportar el diccionario generado para un posterior uso en la aplicación.
- **RF-15 Obtener posiciones del diccionario.** Se podrán obtener todas las posiciones de los personajes dentro del fichero ePub.
- **RF-16 Generar red.** Se podrá generar una red de interacciones teniendo en cuenta el número mínimo de apariciones del personaje en la red, el radio de palabras para generar la interacción y teniendo en cuenta el cambio de capítulo o no.
- **RF-17 Exportar la red generada.** Se podrá exportar la red generada a partir de los parámetros introducidos por el usuario a distintos formados.
- **RF-18 Visualizar red.** Se podrán visualizar la red generada a partir de los parámetros introducidos por el usuario
- **RF-19 Generar informes.** Se podrán obtener distintas métricas acerca de la red generada.
- **RF-20 Internacionalización.** Se podrá utilizar la aplicación en dos de los idiomas más importantes actuales como son el español y el inglés.

Requisitos no funcionales

- **RNF-1 Usabilidad:** La aplicación deberá ser de fácil de usar para nuevos usuarios.
- **RNF-2 Rendimiento:** Se buscará que los costes de tiempo no sean elevados al tratarse de una aplicación web.
- **RNF-3 Mantenibilidad:** El diseño de la aplicación será fácilmente mantenible y extensible.
- **RNF-4 Internacionalización:** La aplicación estará disponible varios idiomas importantes.
- **RNF-5 Seguridad:** La aplicación no mantendrá ficheros con información sensible más tiempo del estrictamente necesario.

B.4. Especificación de requisitos

Actores:

- **Usuario no identificado:** Puede introducir un fichero ePub y cambiar de idioma la aplicación.
- **Usuario identificado:** Puede utilizar la aplicación al completo.

CU - 01	Introducción de archivo ePub	
Descripción	Cualquier usuario puede introducir un archivo ePub para trabajar con él	
Precondiciones	Una carpeta con nombre el identificador de usuario para almacenar información relativa a esa sesión	
Secuencia Normal	Paso	Acción
	1	Si usuario no se ha conectado nunca o se ha perdido su sesión en el servidor se redirige al inicio de la aplicación
	2	Se permite al usuario introducir un archivo ePub
	3	Se comprueba que el archivo introducido es efectivamente un archivo ePub cuando el usuario pulsa el botón “Cargar”
	4	Se guarda el texto del ePub introducido
Excepciones	Paso	Acción
	2	La extensión del archivo no es válida
	3	El fichero no tiene la estructura de los archivos ePub
Postcondiciones	Se carga el fichero y se mueve a la sección diccionarios automáticos	
Frecuencia esperada	Alta	
Importancia	Alta	
Comentarios	Cumple RF-1, RF-2, RF-3	

Tabla B. 1 Introducir ePub

CU - 02	Creación automática de un diccionario de personajes	
Descripción	Un usuario crea un diccionario de personajes a partir de mayúsculas en el ePub subido	
Precondiciones	Haber introducido un archivo ePub en la sesión actual	
Secuencia Normal	Paso	Acción
	1	Desde la sección de diccionarios automáticos pulsar “Crear Diccionario”
	2	Se buscan palabras en el ePub en mayúsculas y se añaden al diccionario de personajes
Excepciones	Paso	Acción
Postcondiciones	Se añaden los personajes encontrados al diccionario actual	
Frecuencia esperada	Baja	
Importancia	Media	
Comentarios	Cumple RF-5	

Tabla B. 2 Crear diccionario personajes

CU - 03	Importación de un diccionario de personajes	
Descripción	Un usuario crea un diccionario de personajes a partir de un fichero csv	
Precondiciones	Haber introducido un archivo ePub en la sesión actual	
Secuencia Normal	Paso	Acción
	1	Desde la sección de diccionarios automáticos pulsar “Importar Diccionario”
	2	Introducir el fichero csv con el diccionario
	3	Pulsar el botón “Cargar”
	4	Se recorre el archivo csv y se añaden los personajes
Excepciones	Paso	Acción
	3	El fichero introducido no tiene la extensión permitida
	4	El formato del archivo csv no es utf-8 y da problemas con algunos caracteres
Postcondiciones	Se añaden los personajes del csv introducido al diccionario actual	
Frecuencia esperada	Media	
Importancia	Media	
Comentarios	Cumple RF-5	

Tabla B. 3 Importar diccionario

CU - 04	Obtención de un diccionario de personajes	
Descripción	Un usuario crea un diccionario de personajes a partir de los personajes que hay en una url específica	
Precondiciones	Haber introducido un archivo ePub en la sesión actual	
Secuencia Normal	Paso	Acción
	1	Desde la sección de diccionarios automáticos pulsar “Obtener Diccionario”
	2	Introducir la url a la wiki con los personajes
	3	Pulsar el botón “Obtener”
	4	Se recorre la dirección en búsqueda de personajes
Excepciones	Paso	Acción
	3	El url introducida no pertenece a la wiki preestablecida
Postcondiciones	Se añaden los personajes de la url introducida al diccionario actual	
Frecuencia esperada	Baja	
Importancia	Baja	
Comentarios	Cumple RF-6	

Tabla B. 4 Obtener diccionario

CU - 05	Vaciar diccionario	
Descripción	Un usuario vacía el diccionario de personajes actual	
Precondiciones	Haber introducido un archivo ePub en la sesión actual	
Secuencia Normal	Paso	Acción
	1	Desde la sección de diccionarios automáticos pulsar “Vaciar Diccionario”
Excepciones	Paso	Acción
Postcondiciones	Se eliminan todos los personajes del diccionario actual	
Frecuencia esperada	Baja	
Importancia	Baja	
Comentarios	Cumple RF-7	

Tabla B. 5 Vaciar diccionario

CU - 06	Añadir Personaje	
Descripción	Un usuario añade un nuevo personaje al diccionario de personajes actual	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Añadir Personaje”
	2	Introducir los valores del nuevo personaje
	3	Pulsar “Añadir”
Excepciones	Paso	Acción
	3	Los valores introducidos no son válidos
Postcondiciones	Se añade el personaje introducido al diccionario	
Frecuencia esperada	Baja	
Importancia	Alta	
Comentarios	Cumple RF-8	

Tabla B. 6 Añadir personaje

CU - 07	Eliminar Personajes	
Descripción	Un usuario elimina personajes del diccionario de personajes actual	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Eliminar Personaje”
	2	Marcar los personajes a eliminar
	3	Pulsar “Eliminar”
Excepciones	Paso	Acción
Postcondiciones	Se eliminan los personajes marcados del diccionario	
Frecuencia esperada	Media	
Importancia	Media	
Comentarios	Cumple RF-9	

Tabla B. 7 Eliminar personaje

CU - 08	Juntar Personajes	
Descripción	Un usuario junta personajes del diccionario de personajes actual	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Juntar Personajes”
	2	Marcar los personajes a juntar o arrastrar un personaje sobre otro
	3	Pulsar “Juntar” en caso de haber marcado los personajes o soltar el personaje arrastrado
Excepciones	Paso	Acción
Postcondiciones	Se eliminan los personajes marcados del diccionario	
Frecuencia esperada	Baja	
Importancia	Media	
Comentarios	Cumple RF-10	

Tabla B. 8 Juntar personajes

CU - 09	Añadir Referencias a Personajes	
Descripción	Un usuario añade una nueva referencia a un personaje del diccionario de personajes actual	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Añadir Referencia”
	2	Introducir los valores de la nueva referencia
	3	Pulsar “Añadir”
Excepciones	Paso	Acción
	3	Los valores introducidos no son válidos
Postcondiciones	Se añade la referencia introducida al personaje del diccionario	
Frecuencia esperada	Media	
Importancia	Alta	
Comentarios	Cumple RF-11	

Tabla B. 9 Añadir nueva referencia

CU - 10	Eliminar Referencias de Personajes	
Descripción	Un usuario elimina las referencias de personajes del diccionario de personajes actual	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Eliminar Referencia”
	2	Marcar las referencias a eliminar
	3	Pulsar “Eliminar”
Excepciones	Paso	Acción
Postcondiciones	Se eliminan las referencias de personajes marcadas del diccionario	
Frecuencia esperada	Baja	
Importancia	Media	
Comentarios	Cumple RF-12	

Tabla B. 10 Eliminar referencias de personajes

CU - 11	Modificar id de Personajes	
Descripción	Un usuario modifica la id de un personaje del diccionario de personajes actual	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Modificar id Personaje”
	2	Introducir los valores para cambiar la id
	3	Pulsar “Cambiar”
Excepciones	Paso	Acción
	3	Los valores introducidos no son válidos
Postcondiciones	Se modifica la id del personaje del diccionario	
Frecuencia esperada	Media	
Importancia	Media	
Comentarios	Cumple RF-13	

Tabla B. 11 Modificar id diccionario

CU - 12	Exportar Diccionario de Personajes	
Descripción	Un usuario puede exportar el diccionario generado para su posterior uso	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Exportar diccionario”
Excepciones	Paso	Acción
Postcondiciones	Se descarga el diccionario en formato csv en el dispositivo del usuario	
Frecuencia esperada	Alta	
Importancia	Alta	
Comentarios	Cumple RF-14	

Tabla B. 12 Exportar diccionario

CU - 13	Obtener Posiciones de Personajes	
Descripción	Un usuario puede obtener las posiciones de los personajes en el texto	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección modificar diccionario Casilla de obtener posiciones marcada	
Secuencia Normal	Paso	Acción
	1	Desde la sección de modificar diccionario pulsar “Siguiente”
	2	Se buscan las posiciones de todas las referencias de los personajes en el ePub
	3	Se unifican las posiciones de todas las referencias de un personaje
Excepciones	Paso	Acción
Postcondiciones	Se muestra al usuario el número de apariciones de cada personaje al usuario y se obtienen las posiciones de cada personaje	
Frecuencia esperada	Alta	
Importancia	Alta	
Comentarios	Cumple RF-15	

Tabla B. 13 Obtener posiciones personajes

CU - 14	Generar Red	
Descripción	Un usuario puede generar una red a partir del diccionario de personajes y unos parámetros limitadores	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección parámetros adicionales	
Secuencia Normal	Paso	Acción
	1	Introducir los parámetros adicionales
	2	Pulsar “Obtener red”
	3	Se genera la red a partir de los parámetros introducidos por el usuario
Excepciones	Paso	Acción
	2	Los datos introducidos no son válidos
Postcondiciones	Se genera la red y se mueve a la sección de visualizar red	
Frecuencia esperada	Alta	
Importancia	Alta	
Comentarios	Cumple RF-16	

Tabla B. 14 Generar red

CU - 15	Exportar la Red Generada	
Descripción	Un usuario puede exportar la red generada para su posterior uso con otras aplicaciones	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección visualizar diccionario	
Secuencia Normal	Paso	Acción
	1	Pulsar el botón “Exportar” con el formato deseado
	2	Se traduce la red de networkx a un fichero con el formato solicitado por el usuario
	3	Se descarga el fichero en el equipo del cliente
Excepciones	Paso	Acción
Postcondiciones	Se descarga la red exportada en el cliente del fichero	
Frecuencia esperada	Media	
Importancia	Media	
Comentarios	Cumple RF-17	

Tabla B. 15 Exportar red

CU - 16	Visualizar la Red Generada	
Descripción	Un usuario puede visualizar la red generada	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección visualizar diccionario	
Secuencia Normal	Paso	Acción
	1	Se envía la red generada a network_styling_with_d3
Excepciones	Paso	Acción
Postcondiciones	Se visualiza la red	
Frecuencia esperada	Alta	
Importancia	Alta	
Comentarios	Cumple RF-18	

Tabla B. 16 Visualizar red

CU - 17	Generar Informes	
Descripción	Un usuario puede generar informes sobre distintas métricas acerca de la red	
Precondiciones	Haber introducido un archivo ePub en la sesión actual Estar en la sección generar informes	
Secuencia Normal	Paso	Acción
	1	El usuario marca las métricas que quiere obtener
	2	Se pulsa “Generar Informe”
	3	Se cambia a la sección visualización de informes
Excepciones	Paso	Acción
	2	Los valores introducidos no son válidos
Postcondiciones	Se visualizan las métricas solicitadas por el usuario	
Frecuencia esperada	Media	
Importancia	Media	
Comentarios	Cumple RF-19	

Tabla B. 17 Generar informes

CU - 18	Internacionalización	
Descripción	Un usuario puede cambiar de idioma	
Precondiciones		
Secuencia Normal	Paso	Acción
	1	Pulsar la bandera del idioma a seleccionar
Excepciones	Paso	Acción
Postcondiciones	Se cambia de idioma	
Frecuencia esperada	Baja	
Importancia	Media	
Comentarios	Cumple RF-20	

Tabla B. 18 Internacionalización

Apéndice C. Especificación de Diseño

C.1. Introducción

Se describen los y argumentan los diseños utilizados durante la realización de la aplicación

C.2. Diseño de datos

Al tratarse de una aplicación web se ha priorizado el objetivo de bajos tiempos de ejecución por encima del bajo uso de memoria.

Datos de sesión

Se guardan en variables de sesión, almacenadas por el propio usuario, los siguientes datos:

- Identificador de usuario.
- Nombre del ePub.
- Idioma para utilizar la aplicación.
- Configuración de la visualización de la red.

Además, en cada sesión por cada usuario conectado se guarda en una carpeta con nombre el identificador de usuario, los ficheros importados o exportados por el usuario durante el uso de la aplicación.

Datos de la base de datos temporal

La base de datos temporal está compuesta por un diccionario y una variable de números enteros.

El diccionario utiliza como claves los identificadores de usuarios guardados en las propias variables de sesión de los usuarios y el objeto modelo asociado a esa sesión.

La variable de números enteros es un contador que incrementa cuando un usuario empieza a utilizar la aplicación, y que se utiliza para informar a cada usuario de su id.

Datos de modelo

El modelo de cada usuario almacena un conjunto de datos que por ser demasiado pesados no se pueden guardar en las variables de sesión, y que por tanto son guardados por el servidor. Los datos son los siguientes:

- Una lista con el texto del libro por cada capítulo.
- Una lista que contiene el número de palabras que tiene cada capítulo.
- Un diccionario de personajes con la id de cada personaje y el objeto personaje asociado a cada personaje.
- El grafo generado.

Datos de personaje

Por cada personaje es necesario tener almacenados una serie de datos para la generación de la red. Se almacenan las siguientes variables por cada personaje:

- Un diccionario con las referencias de ese personaje como id y la posición de estas referencias en cada uno de estos capítulos como valor.
- Un diccionario con las posiciones de todas las posiciones de todas las referencias del personaje, por cada capítulo.
- Un diccionario con el número de referencias que tiene cada referencia del personaje.
- Una variable con el número de apariciones total del personaje.

C.3. Diseño arquitectónico

En esta sección se muestran distintos diagramas de paquetes y de clases.

Diagrama de paquetes

Los paquetes de la aplicación se distribuyen de la siguiente forma:

- **src:** contiene al controlador y la configuración, además del resto de paquetes del proyecto.
- **src/LecturaFicheros:** contiene la lógica para la lectura de ficheros ePub y csv y la de la escritura de ficheros csv.
- **src/Lexers:** contiene los analizadores léxicos de la aplicación.
- **src/Modelo:** contiene la lógica principal de la aplicación.
- **src/PersistenciaSesiones:** contiene la lógica del almacenamiento de los modelos de cada usuario.
- **src/static:** contiene ficheros a utilizar por las plantillas html.
- **src/templates:** contiene las plantillas html.

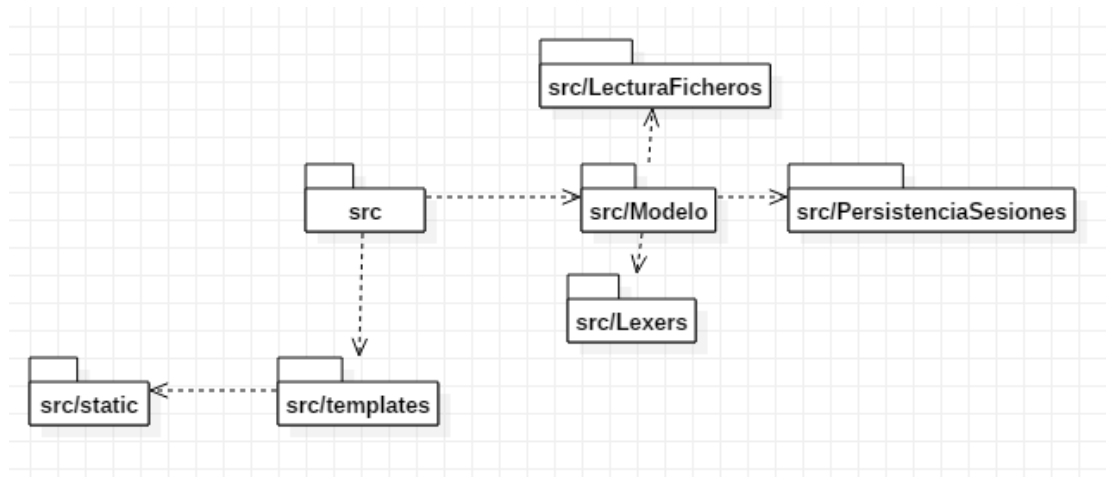


Figura C 1 Diagrama de paquetes

Diagrama de clases

Se muestra cómo se relacionan las clases más importantes del proyecto. En este diagrama no se detallan las variables y los métodos de la clase, debido a que, por su gran cantidad, se perdería claridad. Se dispone del diagrama con los métodos y las variables en doc/DiagramaClasesCompleto.png.

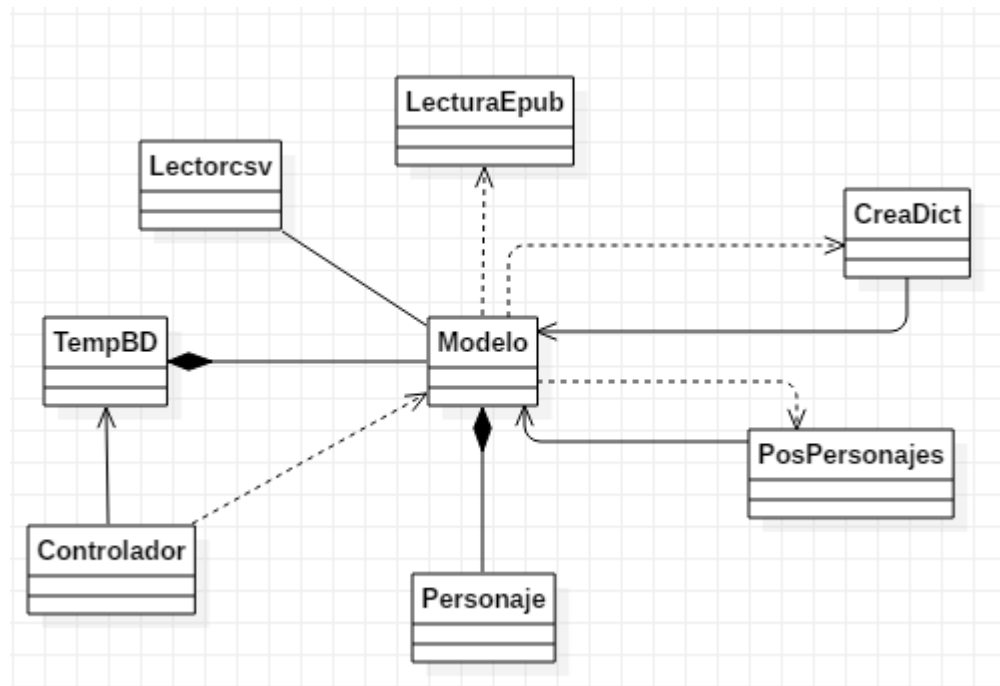


Figura C 2 Diagrama de clases


C.4. Diseño de interfaces

A la hora de diseñar la interfaz, primero se hizo un prototipo de cómo iba a ser la secuencia de trabajo de la aplicación. Para la realización del prototipo se utilizó la herramienta Pencil Project. El prototipado de la interfaz consiguió un ahorro de tiempo, al poder modificar el diseño rápidamente y sin escribir código, antes de empezar a desarrollarla.

El prototipo fue el siguiente:

UBUBOOKNET

INICIO



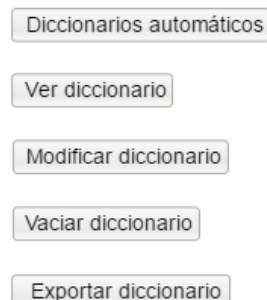
Seleccionar archivo Archivo epub sin seleccionar Cargar

Figura C 3 Prototipo: Inicio 1

UBUBOOKNET

[Inicio](#)

HOME



Diccionarios automáticos

Ver diccionario

Modificar diccionario

Vaciar diccionario

Exportar diccionario

Figura C 4 Prototipo: Menú 2

UBUBOONET

[Inicio](#)

Diccionarios automáticos

☐ Vaciar el diccionario previo

Figura C 5 Prototipo: Diccionarios Automáticos 2.A

UBUBOONET

[Inicio](#)

VER DICCIONARIO

Datos del personaje 1
Datos del personaje 2
Datos del personaje 3

Figura C 6 Prototipo: Ver Diccionario 2.B

UBUBOOKNET[Inicio](#)**MODIFICAR DICCIONARIO**

Anadir personaje

Eliminar personaje

Juntar personajes

Anadir referencia personaje

Eliminar referencia personaje

Datos del personaje 1

Datos del personaje 2

Datos del personaje 3

Volver

Figura C 7 Prototipo: Modificar Diccionario 2.C

Finalmente, este prototipo de interfaz se descarta al considerarse excesivamente complicada para nuevos usuarios, y se adapta para realizar un recorrido lineal y más sencillo de usar por nuevos usuarios.

El diseño actual de la interfaz se muestra en el Anexo de Documentación de Usuario, sección Manual de Usuario.

Apéndice D. Documentación Técnica de Programación

D.1. Introducción

En este anexo se detallan los aspectos relevantes del proyecto para unas correctas futuras tareas de desarrollo y mantenimiento del software por parte del programador.

D.2. Estructura de directorios

Los ficheros que componen el proyecto se encuentran en el siguiente repositorio online <https://github.com/lca0037/GII18.0U-Ububooknet> y tienen la siguiente estructura:

- **/doc/** : Contiene la memoria del proyecto y los diversos anexos.
- **/src/** : Contiene los ficheros para la ejecución del proyecto
- **/src/LecturaFicheros/** : Contiene los ficheros para la importación y exportación de csv y para la lectura del texto del archivo ePub.
- **/src/Lexers/** : Contiene los ficheros para el análisis léxico del texto, tanto para crear un diccionario como para encontrar las posiciones de los personajes.
- **/src/Modelo/** : Contiene tanto la lógica principal de la aplicación como la clase que almacena la información sobre los personajes
- **/src/PersistenciaSesiones/** : Contiene el fichero para el almacenamiento de la información relativa a cada sesión que no puede almacenarse en las variables de sesión de cada usuario, y que se borra cuando el usuario abandona la página.
- **/src/ficheros/** : Directorio para almacenar documentos de exportación e importación relativos a cada usuario.
- **/src/static/** : Contiene archivos relativos a la página web.
- **/src/static/css/** : Contiene los ficheros css encargados de la presentación de la web.
- **/src/static/img/** : Contiene las imágenes a utilizar por la página web.
- **/src/static/js/** : Contiene los ficheros js encargados de la lógica de diversas partes de la web.
- **/src/templates/** : Contiene las plantillas html que utiliza la aplicación.
- **/src/translations/** : Contiene las traducciones que puede realizar babel.

- **/src/Controlador.py** : Fichero que actúa como intermediario entre el modelo y la vista de la aplicación.
- **/src/Config.py** : Archivo de configuración de Flask y de Babel.
- **/tst/** : Directorio que contiene los test de la aplicación.
- **/babel.cfg** : Archivo de configuración de babel.
- **/main.py** : Archivo que ejecuta la aplicación.
- **/requirements.txt** : Archivo que contiene las dependencias a instalar.

D.3. Manual del programador

Instalación

Como se ha comentado en el punto previo, el proyecto se encuentra en <https://github.com/lca0037/GII18.0U-Ububooknet>, pero para la ejecución además de tener el proyecto descargado se necesitan tener instalados los siguientes requisitos:

- Python: versión 3.6
- Flask: versión 1.0
- Flask_Babel: versión 0.12
- numpy: versión 1.15
- matplotlib: 3.0
- ply: versión 3.11
- beautifulsoup4: versión 4.7
- lxml: versión 4.3
- html5lib: versión 1.0
- networkx: versión 2.2
- scipy: versión 1.1

No se necesita ningún requisito adicional para el despliegue en Heroku.

Ejecución en local

Para la ejecución en local se deben instalar los requisitos especificados en el punto previo.

Python se puede descargar e instalar desde el siguiente enlace: <https://www.python.org/downloads/>.

Para la instalación del resto de librerías se puede utilizar el comando:

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

Utilizando el fichero requirements.txt proporcionado en el repositorio.

Para la ejecución de la aplicación se deberá utilizar el siguiente comando:

```
python main.py
```

Utilizando el fichero main.py proporcionado en el repositorio.

```
(base) C:\Users\luism\Desktop\Ubu\TFG\Repositorio\GII18.0U-Ububooknet>python main.py
* Serving Flask app "src.Controlador" (lazy loading)
* Environment: production
  WARNING: Do not use the development server in a production environment.
  Use a production WSGI server instead.
* Debug mode: off
* Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
```

Figura D 1 Ejecución de la aplicación

Despliegue en Heroku

Para el despliegue en Heroku se puede seguir el siguiente [tutorial](#) proporcionado por Heroku para su despliegue mediante línea de comandos.

También se puede desplegar desde la propia página de la aplicación como se va a explicar a continuación.

Los primeros pasos son comunes a ambas formas de desplegar, lo primero que se debe hacer es crear un nuevo repositorio de GitHub donde alojar la aplicación.

Después se deben añadir tres ficheros para la correcta ejecución de la aplicación sobre el servidor:

- Procfile: es un fichero que contiene el comando para ejecutar la aplicación.
- runtime.txt: es un fichero que indica el lenguaje de programación sobre el que está desarrollada la aplicación.
- requirements.txt: es un fichero que indica al servidor las librerías que necesita la aplicación para un correcto funcionamiento.

Una vez tenemos esto, para el despliegue debemos ir a <https://dashboard.heroku.com/apps>, una vez allí pulsaremos el botón “New” que se encuentra arriba a la derecha y seleccionaremos “Create new app”.

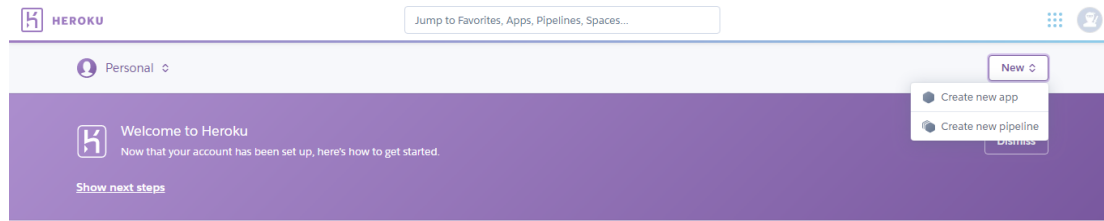


Figura D 2 Crear nueva app heroku

Una vez ahí podremos elegir el nombre que tiene la aplicación y la región donde se quiere alojar el servidor entre Estados Unidos y Europa.

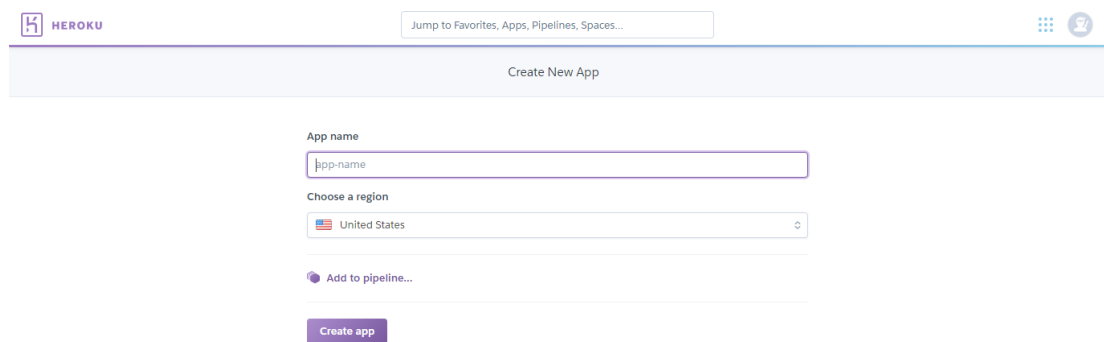


Figura D 3 Información para crear la app

Una vez generada la aplicación, se puede acceder a ella y seleccionar el menú “Deploy”, una vez en él se puede conectar Heroku con la cuenta de GitHub y seleccionar el repositorio desde el que desplegar la aplicación.

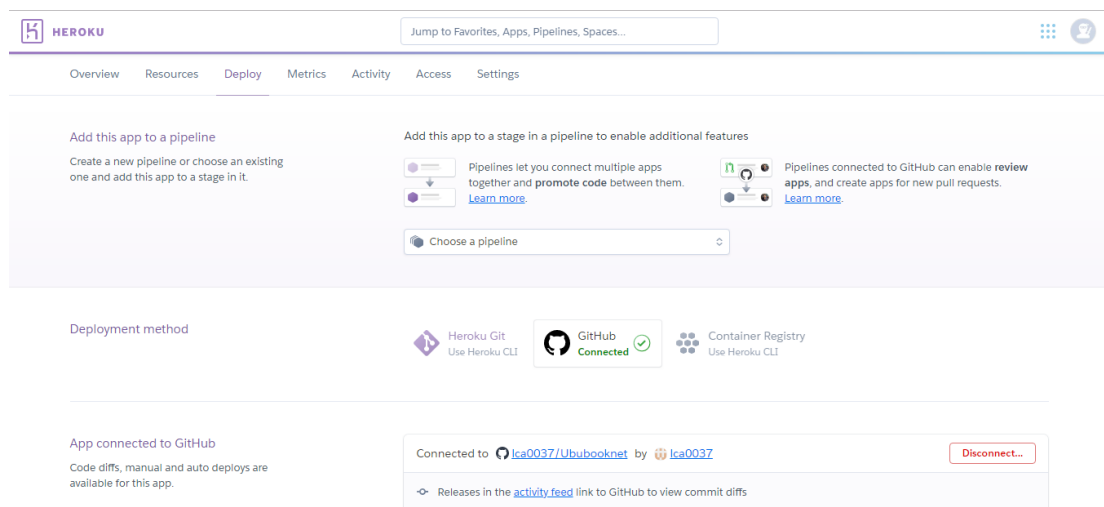


Figura D 4 Vincular con GitHub

Una vez vinculada la cuenta y seleccionado el repositorio se puede decidir si se quieren permitir las actualizaciones automáticas, o si se quiere desplegar manualmente, y la rama del repositorio a desplegar. En caso de no utilizar las actualizaciones automáticas que ofrece Heroku, cada vez que se realice una actualización del repositorio deberemos volver a desplegar la rama para que se actualice también la aplicación web.

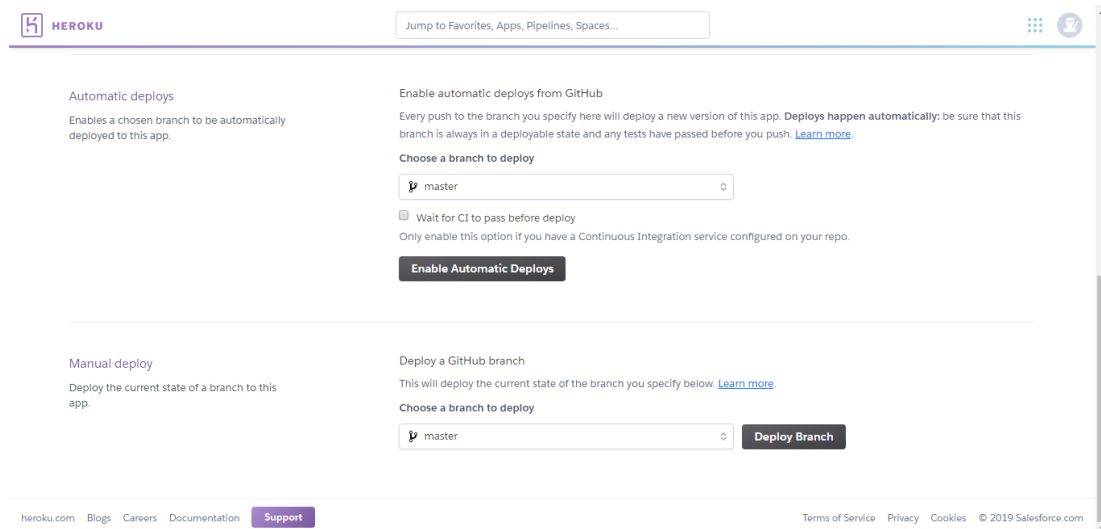


Figura D 5 Selección de la rama

Actualización de los ficheros de internacionalización

Para actualizar los ficheros de internacionalización con nuevas traducciones, aparte de añadir las correspondientes etiquetas en el código (`gettext('')` para código Python y `{{ _('') }}` para html), se deben ejecutar los siguientes comandos, en la copia local del proyecto, sobre el mismo directorio donde se encuentre el fichero `babel.cfg` para actualizar los ficheros de internacionalización:

```
pybabel extract -F babel.cfg -o messages.pot src
pybabel update -i messages.pot -d src/translations
```

Una vez actualizados los ficheros se deberá ir a cada directorio de idioma y abrir el archivo `.po` para actualizar las traducciones.

Nuevos idiomas para la internacionalización

Para la generación de nuevas traducciones se debe ejecutar el siguiente comando en la copia local del proyecto:

```
pybabel init -i messages.pot -d src/translations -l idioma
```

En caso de no generar correctamente los directorios, la ruta al archivo .po y .mo deberá ser /src/translations/idioma/LC_MESSAGES/

También se deberá modificar el fichero /src/Config.py añadiendo el nuevo idioma al diccionario de idiomas.

```
#Idiomas permitidos en internacionalización
LANGUAGES = {
    'en': 'English',
    'es': 'Español'
}
```

Figura D 6 Diccionario idiomas

D.4. Pruebas Unitarias

Se realizan test unitarios a lo largo del desarrollo del proyecto para comprobar el correcto funcionamiento de las funcionalidades implementadas. Estas pruebas se han tenido que actualizar a medida que se modificaban la forma de almacenamiento de los datos.

Para ejecutar las pruebas se deberán quitar los comentarios de las siguientes líneas de código que están actualmente comentadas y comentar las que no lo están.

```
#runner = unittest.TextTestRunner()
#result = runner.run(unittest.makeSuite(testUnitarios.testUnitarios))

Controlador.app.run(threaded=True)
```

Figura D 7 Código a comentar y descomentar

D.5. Scraping para obtener diccionarios de personajes

El scraping se realiza sobre la web <https://www.fandom.com/> y en concreto el apartado de personajes. Aunque actualmente sigue una estructura genérica para todas

las wikis que contiene, esta estructura puede cambiar en cualquier momento y con ello hacer que el scraping que se realiza para obtener el diccionario deje de funcionar correctamente.

Habr  que tener en cuenta estos posibles cambios y actuar en consecuencia, ya sea modificando la funci n que realiza el scraping a la web o buscando una nueva web para obtener el diccionario si es preciso.

A la hora de seleccionar la p gina donde realizar el scraping se busc  que se guardase una consistencia entre todas las p ginas donde sacar los datos, es decir, que aunque se tratase de dos p ginas distintas, la forma de encontrar los datos en el c digo html fuese la misma. Por ejemplo, en la Wikipedia ocurre este problema.

```
<dl>
  <dt> == $0
    <a href="/wiki/Daenerys_Targaryen" title="Daenerys Targaryen">
      Daenerys Targaryen</a>
    </dt>
  ...
```

Figura D 8 Ejemplo Wikipedia Juego de Tronos

```
<ul>
  <li> == $0
    <b>
      <a href="/wiki/Harry_Potter_(personaje)" title=
        "Harry Potter (personaje)">Harry Potter</a>
    </b>
  </li>
</ul>
```

Figura D 9 Ejemplo Wikipedia Harry Potter

Podemos observar c mo las etiquetas que contienen a los personajes son distintas, lo que impide un c digo gen rico para obtener los personajes de ambas p ginas. Es cierto que ambos personajes est n dentro de una etiqueta , pero utilizar esta etiqueta no se considera una opci n debido a que esta etiqueta sirve para insertar enlaces en el c digo html por lo que habr  demasiados falsos positivos.

En la p gina seleccionada, sin embargo, s  que se mantiene una consistencia entre las distintas p ginas.

```
<li class="category-page__member">
  <div class="category-page__member-left">...</div>
  <a href="/wiki/Daenerys_I_Targaryen" title="Daenerys I
    Targaryen" class="category-page__member-link">Daenerys I
    Targaryen</a> == $0
</li>
```

Figura D 10 Ejemplo Fandom Juego de Tronos

```
▼<li class="category-page__member">
  ▶<div class="category-page__member-left">...</div>
  <a href="/wiki/Harry_Potter" title="Harry Potter" class=
    "category-page__member-link">Harry Potter</a> == $0
</li>
```

Figura D 11 Ejemplo Fandom Harry Potter

Como podemos observar, en ambas páginas hay una etiqueta para especificar donde hay un personaje que es `<li class="category-page__member">`. Gracias a esta etiqueta, aunque los nombres de personajes también estén en una etiqueta ``, se pueden distinguir del resto de enlaces, eliminando casi por completo los falsos positivos.

Apéndice E. Documentación de Usuario

E.1. Introducción

En este anexo se detallan los pasos a realizar por el usuario para la correcta ejecución de la aplicación.

E.2. Requisitos de usuarios

Para la correcta ejecución de la aplicación el usuario requiere de un ordenador personal con conexión a internet y cualquiera de estos navegadores:

- Microsoft Edge 12 o superior
- Mozilla Firefox 40 o superior
- Google Chrome 41 o superior
- Safari 9 o superior
- Opera 28 o superior

Aunque en principio se puede acceder desde dispositivos móviles como pueden ser smartphones o tablets, no se garantiza el correcto funcionamiento de la aplicación en estas plataformas.

E.3. Instalación

No requiere ninguna instalación por parte del usuario a parte de lo especificado en el punto previo al tratarse de una aplicación web.

En caso de querer ejecutar la aplicación en local para evitar los problemas de limitación que ofrece el host, se podrá acceder desde la dirección <http://127.0.0.1:5000/>.

E.4. Manual de Usuario

Acceso a la aplicación

Como se ha comentado en el punto previo, se puede acceder a la aplicación en <http://127.0.0.1:5000/> si se está ejecutando en local, o se puede acceder desde <http://ububooknet.herokuapp.com/> al estar la aplicación alojada en un servidor web, aunque se deberá tener en cuenta que se producen timeouts al ser un servidor gratuito.

Características globales - Navegador

Durante todo el uso de la aplicación, en caso de cerrar la pestaña de la aplicación, o de usar los botones indicados en la imagen puesta a continuación se perderá la sesión y con ello los datos generados.



Figura E 1 Botones del navegador

Características globales - Menú

El usuario, esté en la parte que esté de la aplicación dispondrá de un menú donde verá su ubicación en el workflow de la aplicación y los pasos ya realizados más los que le quedan por hacer, pudiendo hacer click en los pasos ya realizados para volver a esa parte de la ejecución, dos banderas donde poder cambiar el idioma de la aplicación simplemente haciendo click en ellas, el nombre de la aplicación a la izquierda, la parte en la que está en el centro y un enlaces al inicio de la aplicación y al repositorio de la misma, y un enlace de Acerca De para explicar las librerías y herramientas utilizadas.

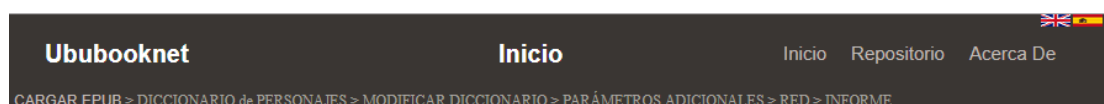


Figura E 2 Menú de la aplicación

Página de inicio

Al inicio de la aplicación se le solicitará al usuario que introduzca un fichero .epub y se comprobará que el contenido del fichero introducido efectivamente se corresponde con el que un fichero .epub debiera tener.

Al estar los ePubs en un formato basado en zip, como se ha explicado en los conceptos teóricos de la memoria, también se permite al usuario subir archivos .zip.



Figura E 3 Interfaz de inicio

Para seleccionar el archivo ePub, tendrá un botón de seleccionar archivo que le abrirá el explorador de archivos y ahí podrá seleccionar el archivo a subir, una vez seleccionado podrá cargarlo para avanzar a la siguiente fase de la aplicación

Diccionarios automáticos

Una vez el usuario haya introducido su fichero ePub, se le permitirá generar un diccionario automático de nombres de diversas formas o vaciar completamente su diccionario de personajes.

A la hora de generar su diccionario de personajes dispone de 4 opciones:

- La primera es pulsar el botón de crear diccionario que buscará en el ePub introducido todas las palabras en mayúsculas.
- La segunda es importar un diccionario, que le llevará a una ventana donde poder introducir su diccionario de personajes, que es la opción que se recomienda.
- La tercera es obtener un diccionario, que también le llevará a otra ventana, pero en esta le permitirá introducir una url de una página específica que será explicada posteriormente.
- La cuarta y última opción es pulsar al botón siguiente sin haber realizado ninguna de las acciones anteriores para generar el diccionario desde la propia página.

Ninguna de estas opciones elimina los personajes introducidos anteriormente.

El usuario además dispone de un botón para volver a la carga del ePub.

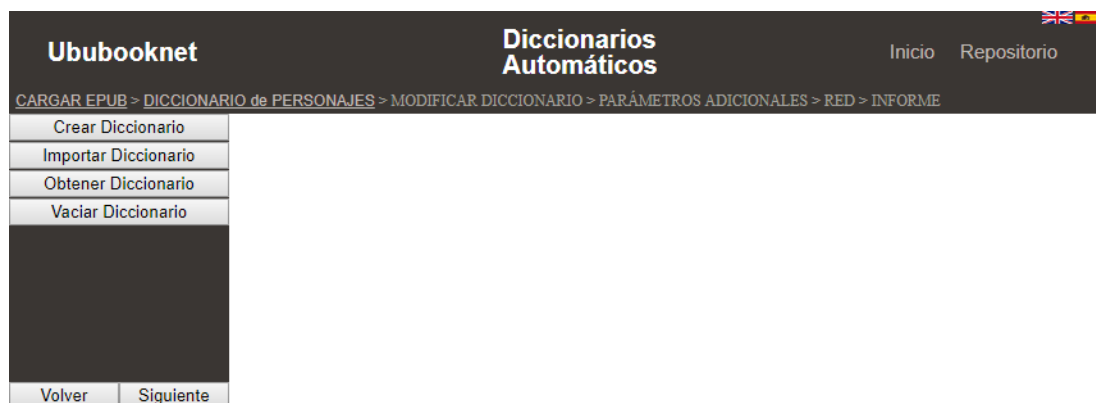


Figura E 4 Interfaz de diccionarios automáticos

Importar diccionario

En esta ventana el usuario podrá introducir un fichero csv de la misma forma que introduce un fichero ePub

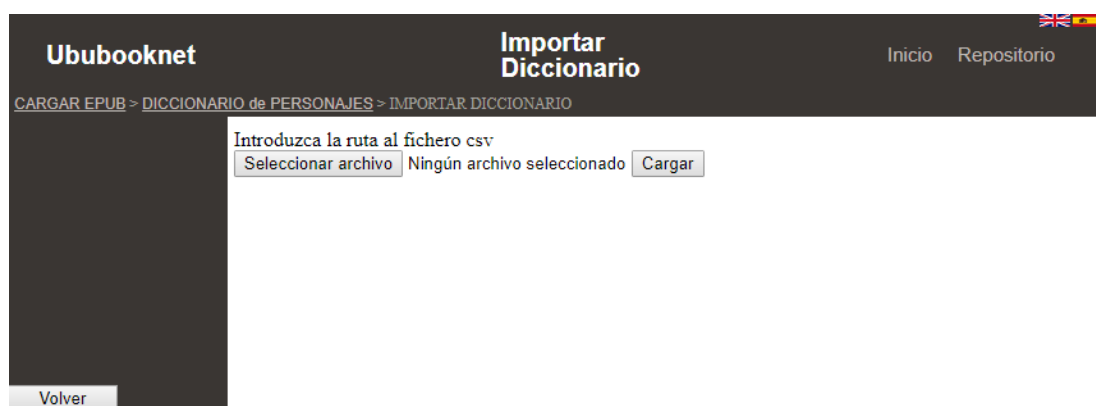


Figura E 5 Interfaz de importar diccionario

El formato del archivo csv a introducir deberá ser utf-8 para un correcto funcionamiento de la aplicación y seguir la siguiente estructura:

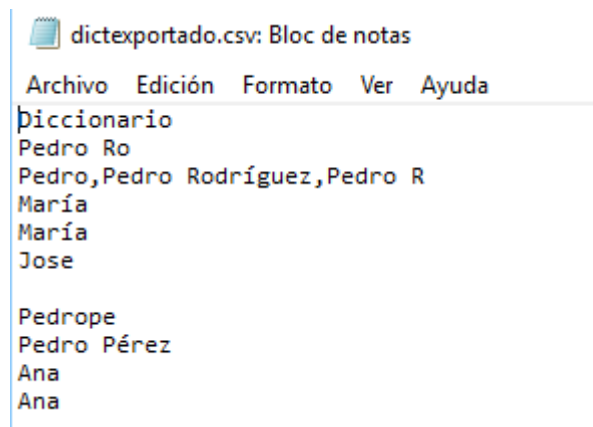


Figura E 6 Ejemplo de archivo csv de personajes

La primera fila contendrá texto que será ignorado a la hora de leer el diccionario, después, las filas se cogerán por pares.

En la primera fila del par de filas vendrá un único nombre que se utilizará como id del personaje y como su primer nombre, y en segunda fila del par, vendrán el resto de los nombres con los que se referencia al personaje separados por comas.

La segunda fila puede estar vacía en el caso se que solo se quiera añadir el mismo nombre que la propia id, y también se puede añadir la propia id a la segunda fila sin que por ello se duplique el nombre.

Obtener diccionario

A la hora de obtener un diccionario mediante scraping, el usuario deberá de introducir el enlace de la wiki en el recuadro de texto y pulsar el botón obtener.

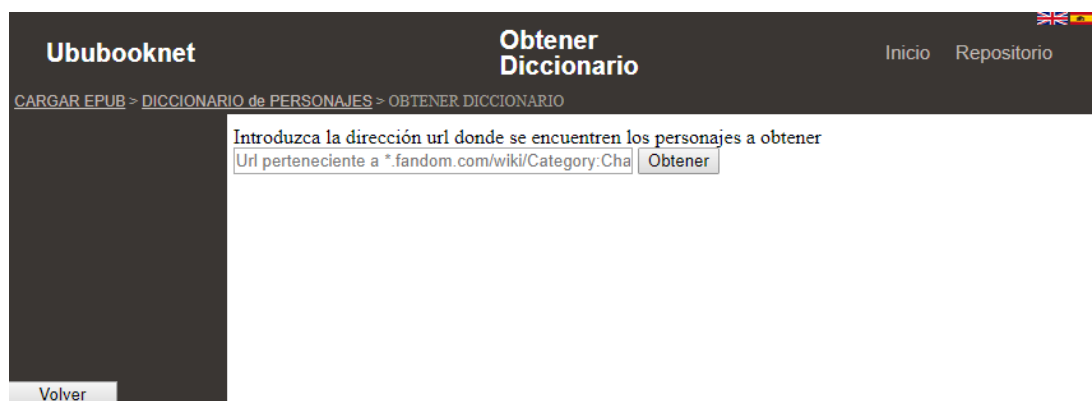


Figura E 7 Interfaz de obtener diccionario

En enlace a la página donde realizar el scraping debe de pertenecer a <https://www.fandom.com/> y en concreto, para un correcto resultado a la hora de obtener los personajes, se debe de estar en la wiki del libro en cuestión en el apartado de personajes, por ejemplo, para obtener los personajes de Juego de Tronos se podría usar este enlace <https://iceandfire.fandom.com/wiki/Category:Characters>.

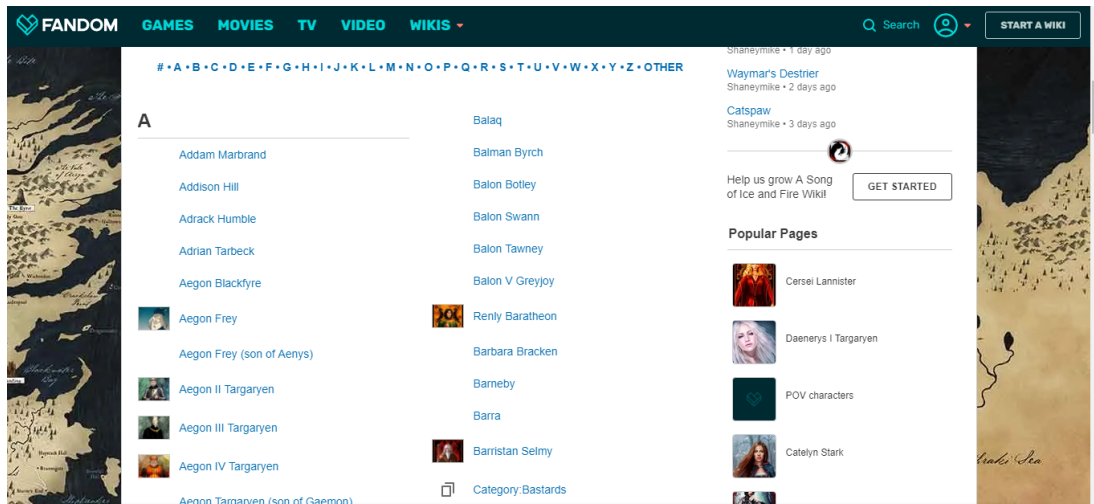


Figura E 8 Vista de la página donde realizar scraping

Los personajes que se añadirán al diccionario serán los pertenecientes a esa url, es decir, en caso de necesitar personajes en otras urls, el usuario tendrá que introducir la url, dar al botón obtener, y seguir este proceso hasta que obtenga los personajes que desee.

Modificar diccionario

Una vez el usuario esté en esta parte de la aplicación, podrá editar su diccionario de personajes de diversas formas:

- Añadiendo nuevos personajes
- Eliminando personajes ya existentes
- Juntando dos o más personajes en uno solo
- Añadiendo referencias a personajes
- Eliminando referencias a personajes
- Modificando la id de los personajes

El usuario además podrá exportar el diccionario generado para futuros usos en la aplicación.

Ububooknet **Añadir Personaje** Inicio Repositorio

CARGAR EPUB > DICCIONARIO de PERSONAJES > MODIFICAR DICCIONARIO > AÑADIR PERSONAJE

introduzca la id del nuevo personaje

ID Personaje

introduzca el nombre del nuevo personaje

Nombre Personaje

Añadir

ID Personaje	Referencias personaje	Número Apariciones
Harry	Harry	N/A
Potter	Potter	N/A
Ron	Ron	N/A
Hermione	Hermione	N/A
Dumbledore	Dumbledore	N/A
Voldemort	Voldemort	N/A

Volver

Figura E 10 Interfaz de añadir personaje

Eliminar personaje

El usuario dispone de una casilla por cada uno de los personajes. Para eliminar personajes, basta con marcar las casillas de todos los personajes a eliminar y después pulsar el botón Eliminar.

Ububooknet **Eliminar Personaje** Inicio Repositorio

CARGAR EPUB > DICCIONARIO de PERSONAJES > MODIFICAR DICCIONARIO > ELIMINAR PERSONAJE

Marque los personajes a eliminar

Eliminar

ID Personaje	Referencias personaje	Número Apariciones
<input type="checkbox"/> Harry	Harry	N/A
<input checked="" type="checkbox"/> Potter	Potter	N/A
<input type="checkbox"/> Ron	Ron	N/A
<input checked="" type="checkbox"/> Hermione	Hermione	N/A
<input type="checkbox"/> Dumbledore	Dumbledore	N/A
<input checked="" type="checkbox"/> Voldemort	Voldemort	N/A

Volver

Figura E 11 Interfaz de eliminar personaje

Juntar personajes

El usuario dispone de una casilla por cada uno de los personajes. Para juntar personajes, basta con marcar las casillas de todos los personajes a juntar y después pulsar el botón Juntar.

Adicionalmente el usuario podrá juntar personajes arrastrando uno encima del otro.

En caso de utilizar las casillas, la id que se mantendrá al juntar los personajes será la situada más arriba en la lista de personajes, y en caso de arrastrar un personaje sobre otro, se mantendrá la del personaje sobre el que se ha soltado el otro.

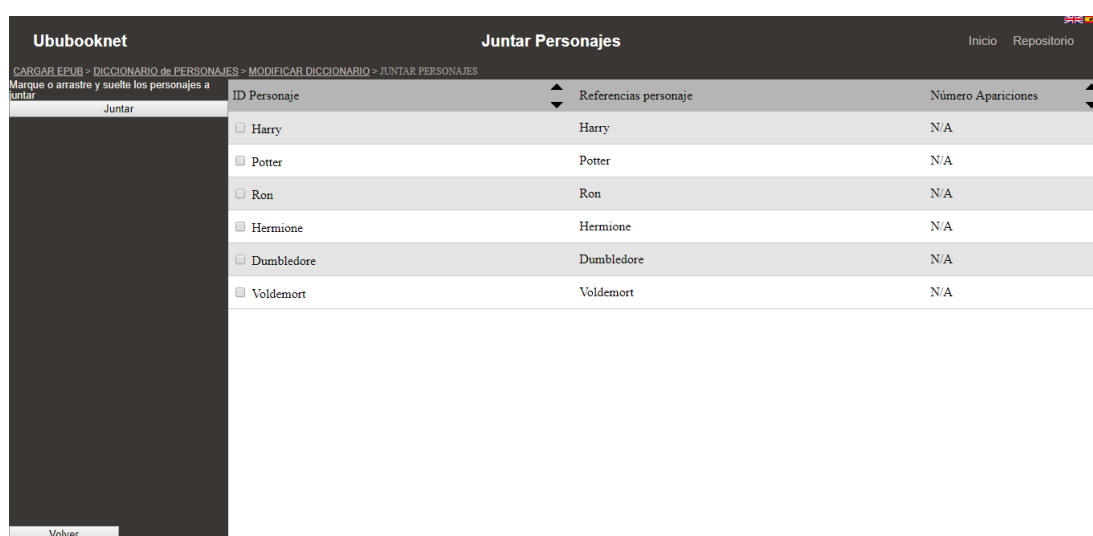


Figura E 12 Interfaz de juntar personajes

Añadir referencia a personaje

El usuario dispone de dos recuadros de texto, el primero para introducir la id del personaje al que añadir una referencia y el segundo para introducir la nueva referencia.

Una vez el usuario haya introducido unos valores a este formulario, podrá pulsar el botón de Añadir, que añadirá la nueva al personaje indicado en caso de que se cumplan los requisitos. En caso de no cumplirse avisará al usuario de cuál ha sido el problema por el que no se ha añadido la nueva referencia.

Ububooknet Inicio Repositorio

CARGAR EPUB > DICCIONARIO de PERSONAJES > MODIFICAR DICCIONARIO > AÑADIR REFERENCIA a PERSONAJE

Añadir Referencia A Personaje

Introduzca la id del personaje a añadir la referencia

ID Personaje	Referencias personaje	Número Apariciones
Harry	Harry	N/A
Potter	Potter	N/A
Ron	Ron	N/A
Hermione	Hermione	N/A
Dumbledore	Dumbledore	N/A
Voldemort	Voldemort	N/A

Volver

Figura E 13 Interfaz añadir referencia

Eliminar referencias de personajes

El usuario dispone de una casilla por cada una de las referencias. Para eliminar referencias, basta con marcar las casillas de todas las referencias a eliminar y después pulsar el botón Eliminar.

En caso de eliminar la única referencia de un personaje, también se eliminará al personaje.

Ububooknet Inicio Repositorio

CARGAR EPUB > DICCIONARIO de PERSONAJES > MODIFICAR DICCIONARIO > ELIMINAR REFERENCIA PERSONAJE

Eliminar Referencia De Personaje

Marque las referencias de personajes a eliminar

ID Personaje	Referencias personaje	Número Apariciones
Harry	<input type="checkbox"/> Harry	N/A
Potter	<input type="checkbox"/> Potter	N/A
Ron	<input type="checkbox"/> Ron	N/A
Hermione	<input type="checkbox"/> Hermione	N/A
Dumbledore	<input type="checkbox"/> Dumbledore	N/A
Voldemort	<input type="checkbox"/> Voldemort	N/A

Volver

Figura E 14 Interfaz de eliminar referencia

Modificar id de un personaje

El usuario dispone de dos recuadros de texto, el primero para introducir la id actual del personaje y el segundo para introducir la nueva id para el personaje.

Una vez el usuario haya introducido unos valores a este formulario, podrá pulsar el botón de Cambiar, que modificará la id al personaje indicado en caso de que se cumplan los requisitos. En caso de no cumplirse avisará al usuario de cuál ha sido el problema por el que no se ha modificado la id.

The screenshot shows the 'Modificar Id Personaje' interface. On the left is a sidebar with the following links: 'CARGAR EPUB > DICCIONARIO de PERSONAJES > MODIFICAR DICCIONARIO > MODIFICAR ID PERSONAJE', 'Introduzca la id actual del personaje', 'Introduzca la nueva id', 'Nueva Id', 'Cambiar', and 'Volver'. The main area contains a table with the following data:

ID Personaje	Referencias personaje	Número Apariciones
Harry	Harry	N/A
Potter	Potter	N/A
Ron	Ron	N/A
Hermione	Hermione	N/A
Dumbledore	Dumbledore	N/A
Voldemort	Voldemort	N/A

Figura E 15 Interfaz modificar id personaje

Parámetros adicionales

El usuario dispone de dos recuadros de texto, el primero para introducir el número mínimo de apariciones, de forma que si no cumple el mínimo no aparezca representado en la red, y la distancia de palabras que puede haber como máximo entre dos posiciones de distintos personajes para considerar que hay una interacción entre ambos.

Además, se dispone de una casilla que marcada tiene en cuenta los capítulos a la hora de las interacciones entre personajes y desmarcada no.

Generación de informes

En esta ventana se disponen de casillas donde marcar información a obtener, en el caso de los k-clique para comunidades además se pide al usuario que especifique el valor de k. En caso de introducir valores correctos, al pulsar generar informe se llevará al usuario a una ventana donde poder ver las métricas generadas.

Al igual que ocurría con la obtención de personajes, se recomienda la utilización de una red pequeña a la hora de obtener las comunidades por Girvan-Newman y de detectar los roles cuando no se ejecute la aplicación en local, debido a que, al ser operaciones muy costosas en tiempo, estas pueden derivar en un timeout.

Se agrupan las medidas que se pueden obtener en la generación de informes de la siguiente forma:

- Medidas básicas
 - Obtener el número de nodos
 - Obtener el número de enlaces
 - Obtener el número de interacciones
- Medidas de grado
 - Obtener el grado de la red sin el peso
 - Obtener el grado de la red con el peso
 - Obtener la distribución de grado de la red sin el peso
 - Obtener la distribución de grado de la red con el peso
 - Obtener la densidad de la red
- Medidas de distancia
 - Obtener los componentes conectados en la red
 - Obtener la excentricidad de la red
 - Obtener el diámetro de la red
 - Obtener el radio de la red
 - Obtener la distancia media de la red
- Medidas de centralidad
 - Obtener la centralidad de grado
 - Obtener la centralidad de cercanía
 - Obtener la centralidad de intermediación
 - Obtener la centralidad de camino aleatorio
 - Obtener la centralidad de valor propio
 - Obtener la centralidad mediante pagerank
- Obtención de grupos y comunidades
 - Obtener comunidades K-clique
 - Obtener comunidades mediante Girvan-Newman
- Detección de roles de la red
 - Detectar roles en comunidades Girvan-Newman

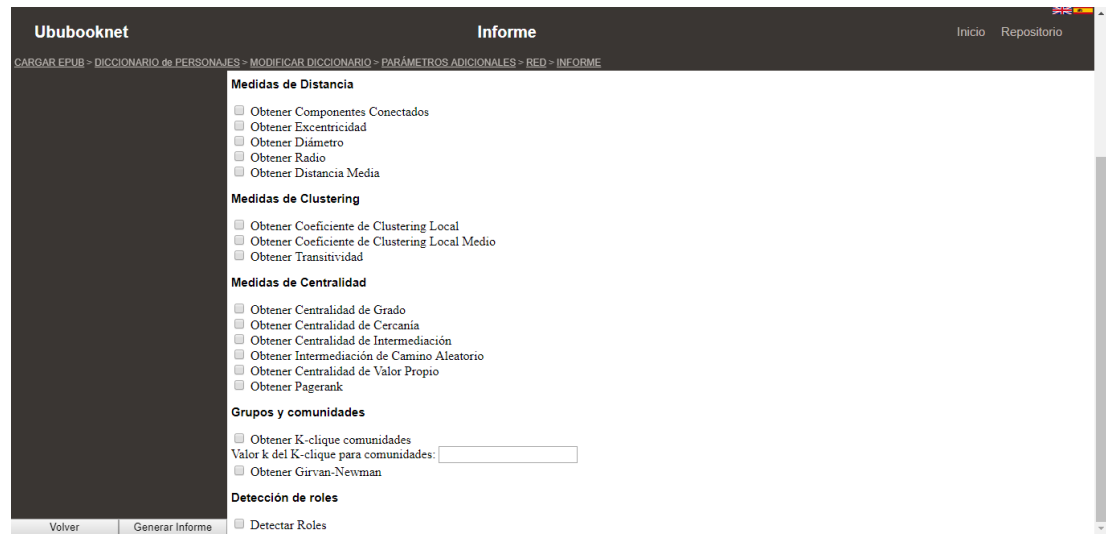


Figura E 18 Interfaz generación de informes