



**UNIVERSIDAD
DE GRANADA**

TRABAJO FIN DE MASTER
INGENIERÍA DE TELECOMUNICACIÓN

Herramienta para la importación de CVN a la plataforma SICA

Autor

Juan Carlos Angulo Santos

Directores

Jorge Navarro Ortiz



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS INFORMÁTICA Y DE
TELECOMUNICACIÓN

Granada, Septiembre de 2018

Herramienta para la importación de CVN a la plataforma SICA

Autor

Juan Carlos Angulo Santos

Director

Jorge Navarro Ortiz

Herramienta para la importación de CVN a la plataforma SICA

Juan Carlos Angulo Santos

Palabras clave: CVN, CV, SICA, Fecyt, Curriculum Vitae Normalizado, Web, Análisis de tráfico, Importar, Investigadores

Resumen

En la actualidad, los investigadores deben incluir su CV (Currículum Vitae) en diferentes plataformas y formatos para distintas convocatorias y entidades. Así, por ejemplo, el Ministerio de Economía, Industria y Competitividad utiliza un Currículum Vitae Normalizado (CVN) para muchas de sus convocatorias, contando con un editor para que cada investigador pueda gestionar y generar dicho CVN. Por otro lado, en el caso de Andalucía, los investigadores se incluyen en grupos de investigación del Sistema de Información Científico Andaluz (SICA). Esta plataforma también permite la gestión del CV de los investigadores, a través de la interacción con el usuario e indagando en otras fuentes de información científica.

Sin embargo, aunque la plataforma SICA permite la exportación al formato CVN, no se permite el paso contrario, es decir, su importación. Este Trabajo de Fin de Máster pretende abordar esta carencia, permitiendo así que un investigador que tenga generado un CVN (sea con el editor del Ministerio de Economía, Industria y Competitividad o de cualquier otra forma) pueda importar los datos de su CVN a SICA de forma automatizada, en lugar de introducir nuevamente todos los elementos que incluye dicho documento.

Tool for the import of CVN to the SICA platform

Juan Carlos Angulo Santos

Keywords: CVN, CV, SICA, Fecyt, Standardized Curriculum Vitae, Web, Traffic analysis, Import, Researchers

Abstract

Currently, researchers must include their CV (Curriculum Vitae) in different platforms and formats for different calls and entities. So, for example, the Spanish Ministry of Economy, Industry and Competitiveness uses a Standardized Curriculum Vitae (CVN) for many of its calls, with an editor so that each researcher can manage and generate such CVN. On the other hand, in the case of Andalusia, researchers are included in research groups of the Andalusian Scientific Information System (SICA). This platform also includes the management of researchers' CVs, through interaction with the user and querying other sources of scientific information

However, although the SICA platform allows export to the format CVN, the complementary operation, namely its importation, is not allowed. This Master's Thesis aims to address this weakness, thus enabling a researcher who has generated a CVN (either with the editor of the Spanish Ministry of Economy, Industry and Competitiveness or in any other way) may import data from its CVN to SICA, instead of introducing again all the elements included in such a document.

Yo, **Juan Carlos Angulo Santos**, alumno de la titulación del Máster Universitario en Ingeniería de Telecomunicación de la **Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación de la Universidad de Granada**, con DNI , autorizo la ubicación de la siguiente copia de mi Trabajo Fin de Master en la biblioteca del centro para que pueda ser consultada por las personas que lo deseen.

Fdo: Juan Carlos Angulo Santos

Granada a septiembre de 2018.

D. **Jorge Navarro Ortiz**, Profesor del Área de Ingeniería Telemática del Departamento de Teoría de la Señal, Telemática y Comunicaciones de la Universidad de Granada.

Informa:

Que el presente trabajo, titulado ***Herramienta para la importación de CVN a la plataforma SICA***, ha sido realizado bajo su supervisión por **Juan Carlos Angulo Santos**, y autorizo la defensa de dicho trabajo ante el tribunal que corresponda.

Y para que conste, expide y firma el presente informe en Granada a septiembre de 2018 .

El director:

Jorge Navarro Ortiz

Índice general

Índice de figuras	12
Índice de tablas	17
1. Introducción.	1
2. Estado del arte.	5
3. Requisitos.	11
4. Planificación y presupuesto.	13
4.1. Planificación	13
4.2. Presupuesto	16
4.2.1. Recursos humanos	16
4.2.2. Recursos hardware	17
4.2.3. Recursos software	18
4.2.4. Presupuesto final	18
5. Currículum Vitae Normalizado.	19
5.1. El estándar CVN	19
5.2. La norma CVN	21
6. Diseño e implementación.	33
6.1. Diseño	33
6.1.1. Lenguajes utilizados	33
6.1.2. Man in the middle proxy (mitmproxy)	34
6.2. Implementación	38
6.2.1. Análisis del formato CVN	38
6.2.2. Análisis de tráfico en SICA	42
6.2.3. Desarrollo de la herramienta	50
7. Fase de pruebas.	79

8. Conclusiones y líneas futuras.	89
8.1. Conclusiones	89
8.2. Líneas futuras	90

Índice de figuras

1.1.	Opción de descarga de CVN en SICA	2
1.2.	Opciones de gestión curricular en SICA	3
2.1.	Extracto del listado de instituciones que permiten exportar e importar CVN	6
2.2.	Importación de distintos formatos de CV a FECYT	7
2.3.	Plataforma Europass	8
2.4.	Importación de trabajos en ORCID	9
2.5.	Importación en diferentes formatos a EndNote	10
4.1.	Diagrama de Gantt	16
5.1.	Ejemplo de formato CVN-XML	22
5.2.	Nodos principales del CVN-XML	24
5.3.	Campos para el nodo <i>Version</i>	25
5.4.	Campos para el nodo <i>Agent</i> (I)	27
5.5.	Campos para el nodo <i>Agent</i> (II)	28
5.6.	Campos referentes al nodo <i>CvnItem</i> para una aportación a congreso	31
6.1.	Interfaz de mitmweb	35
6.2.	Funcionamiento de mitmproxy explícito	35
6.3.	Conexión y creación de certificados utilizando HTTPS.	37
6.4.	Interfaz de mitmproxy web.	37
6.5.	Ejemplo de CVN-PDF abierto con visualizador de PDF	39
6.6.	Ejemplo de CVN-PDF abierto con editor de textos	40
6.7.	Código Base64 incluido en el CVN-PDF	40
6.8.	Identificador del tipo de elemento en el CV (CNPk).	41
6.9.	Ejemplo de nodos de información en CVN-XML	42
6.10.	Página de inicio de la plataforma SICA	42
6.11.	Página de añadir elemento de categoría Otros en SICA	43
6.12.	Página final para añadir elementos de categoría Otros en SICA	44
6.13.	Diagrama de pasos a seguir para añadir elemento a SICA	44
6.14.	Ejemplo de petición HTTPS en SICA visualizado en mitmweb	45
6.15.	Ejemplo de respuesta desde SICA visualizada en mitmweb	46
6.16.	Campos enviados en el login de SICA	47

6.17. Diagrama que muestra el análisis del tráfico con mitmweb en SICA	47
6.18. Campos del tipo form:j_id	48
6.19. Diagrama que muestra como encontrar campos de tipo dinámico	49
6.20. Campos de formulario con petición AJAX	50
6.21. Página de inicio de sesión en la herramienta	52
6.22. Página de inicio de sesión adaptable en dispositivo móvil	53
6.23. Información sobre el usuario que ha iniciado sesión	53
6.24. Mensaje de error en la página de inicio de sesión	54
6.25. Página para introducir fichero CVN-PDF o CVN-XML	55
6.26. Ventana para la selección de fichero	56
6.27. Error de CVN-PDF no válido en la página de importación	56
6.28. Página de pre-selección para la resolución de conflictos en el CVN y SICA	58
6.29. Página de selección de elementos a importar a SICA	59
6.30. Campos de un elemento en la página de selección	60
6.31. Final de la página de selección de elementos a importar a SICA	61
6.32. Elemento con campos obligatorios sin rellenar	61
6.33. Listado de elementos que no pudieron ser importados a SICA	62
6.34. JSON enviado para el procesamiento de importación a SICA	62
6.35. Diagrama de flujo del proceso de importación en el lado del cliente	63
6.36. Vector con todos los índices con códigos de identificación	64
6.37. Diagrama en el que se muestra el vector principal	65
6.38. Diagrama en el que se muestra el vector del código de identificación	66
6.39. Variables con todas las secciones y subsecciones encontradas en el editor de Fecyt	66
6.40. Índices y valores para el índice del código de elemento	67
6.41. Diagrama en el que se muestra el vector de <i>categorias</i>	68
6.42. Índices y valores para el índice de <i>categorias</i>	68
6.43. Diagrama en el que se muestra el vector de <i>tipos_item_sica</i>	69
6.44. Índices y valores para el índice de <i>tipos_item_sica</i>	70
6.45. Identificadores de tipo <i>item</i> en SICA	70
6.46. Diagrama en el que se muestra el vector de <i>campos_mostrar</i>	71
6.47. Índices y valores para el índice de <i>campos_mostrar</i>	72
6.48. Diagrama en el que se muestra el vector de <i>pasos</i>	74
6.49. Índices y valores para el índice de <i>pasos</i>	75
7.1. Configuración de área local	80
7.2. Mitmweb con filtrado por SICA	81
7.3. Categoría del <i>item</i> a escoger en SICA	81
7.4. Tipo y subtipo del <i>item</i> a escoger en SICA	81
7.5. Campos rellenados para un <i>item</i> de SICA	82
7.6. Nombre de un campo en SICA	82
7.7. Campo añadido al vector <i>campos_mostrar</i>	82
7.8. <i>Item</i> añadido a SICA	83
7.9. Pasos añadidos a la herramienta para importar un <i>item</i> a SICA	83
7.10. Exportación de CVN-XML desde SICA	84

7.11. Explotación de CVN-XML del <i>item</i> añadido a SICA	84
7.12. Identificación de <i>item</i> y los valores de los campos	85
7.13. CVN-XML añadido en la herramienta	85
7.14. Selección de <i>item</i> añadido a la herramienta	86
7.15. Momento de importación en la herramienta	87
7.16. Importación correcta del <i>item</i> seleccionado	87
7.17. <i>Item</i> importado a SICA correctamente	88

Lista de Tablas

4.1.	Estudio de costes para los recursos humanos.	17
4.2.	Presupuesto final del proyecto.	18
6.1.	Páginas existentes en la herramienta.	51

Capítulo 1

Introducción.

En este capítulo se realizará una introducción referente a este Trabajo de Fin de Máster, incluyendo la contextualización, motivación y objetivos del mismo.

En la actualidad, los investigadores deben incluir su Currículum Vitae (CV) en diferentes plataformas y formatos para distintas convocatorias y entidades. Así, por ejemplo, el Ministerio de Economía, Industria y Competitividad utiliza un formato de CV conocido como Currículum Vitae Normalizado (CVN) [1], perteneciente a la Fundación Española para la Ciencia y la Tecnología (FECYT) [2], para muchas de sus convocatorias. Esta entidad proporciona un editor para que cada investigador gestione y genere dicho CVN a través de una dirección Web.

Como se detallará más adelante, el CVN, es una norma estándar que define un mismo formato de presentación de los datos curriculares de los investigadores y permite que estos puedan incluirse en las diferentes bases de datos de las instituciones. En definitiva, sirve para que los investigadores puedan presentar su CV en un formato digital único y normalizado, un formato que les permite actualizar de forma automática todos sus datos, sin tener que volver a introducirlo de nuevo.

En el caso de Andalucía, los investigadores se incluyen en Grupos de Investigación del Sistema de Información Científico de Andalucía (SICA) [3], una aplicación de la Secretaría General de Universidades, Investigación y Tecnología adscrito a la Agencia Andaluza de Evaluación de la Calidad y la Acreditación Universitaria, cuyo objetivo principal es recopilar y ofrecer información sobre la actividad científica e investigadora de las Universidades Andaluzas [4].

Esta aplicación que sirve de apoyo curricular para los investigadores de las diferentes universidades de Andalucía se divide en:

- **Investig-An:** Gestor curricular de la producción y actividad científica de los agentes del Sistema Andaluz del Conocimiento estructurado según el estándar CVN.

- **Conect-An:** Gestor de redes sociales basadas en el conocimiento para establecer espacios de colaboración entre los agentes del Sistema Andaluz del Conocimiento.
- **Export-An:** Extractor de actividad y producción científica asociada a dominios institucionales.
- **Agrup-An:** Gestor de grupos del Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI).

El presente proyecto se basará en la sección *Investig-An*, en la cual se realiza la gestión curricular de la actividad científica basándose en el estándar CVN. SICA permite la exportación de los datos introducidos en la plataforma al formato CVN, como se observa en la figura 1.1, en la cual se muestra que existe la opción de exportar en formato CVN-XML o PDF-FECYT (CVN-PDF), únicos formatos relacionados con CVN en SICA.

The screenshot shows the SICA platform interface. At the top, there is a search bar with a 'Buscar' button and a link to 'Búsqueda Avanzada'. Below this is a table with columns: Tipo, Título, Rol, Fecha, and a checkbox. The table contains two rows of data. Below the table, there is a section titled 'Criterios de exportación' with two radio buttons: 'Todas las fechas' (selected) and 'Intervalo personalizado'. Below this, there is a label 'Seleccione un formato de exportación' and a dropdown menu. The dropdown menu is open, showing the following options: CVN-XML, PDF-FECYT (highlighted), Currículum abreviado (CVA) - FECYT, Currículum SICA2 (odt), Currículum SICA2 (doc), and Currículum SICA2 (pdf). To the right of the dropdown menu is a 'Exportar' button.

Tipo	Título	Rol	Fecha	
Situación profesional actual	Profesor Contratado Doctor	Docente	01/02/2018	<input checked="" type="checkbox"/>
Artículo	A Simple Model for Predicting ...	Autor/a	2018	<input checked="" type="checkbox"/>

Criterios de exportación

☒ Todas las fechas ☐ Intervalo personalizado

Seleccione un formato de exportación

- CVN-XML
- PDF-FECYT
- Currículum abreviado (CVA) - FECYT
- Currículum SICA2 (odt)
- Currículum SICA2 (doc)
- Currículum SICA2 (pdf)

Exportar

Figura 1.1: Opción de descarga de CVN en SICA

Sin embargo, en SICA no existe ninguna opción para la importación de ficheros, incluido el CVN, lo que resulta en un esfuerzo adicional para los investigadores que ya posean un CVN, el volver a introducir los mismos elementos en SICA. La única opción para introducir elementos en SICA es a través de la interacción del usuario con la plataforma, añadiéndolos uno por uno. En la figura 1.2 se pueden observar las distintas opciones que provee SICA para la gestión curricular en su plataforma:

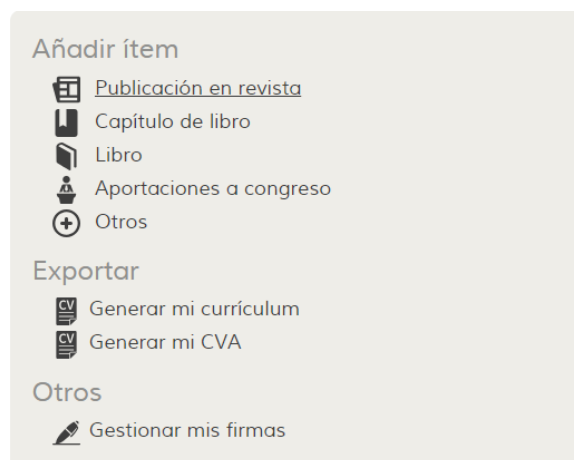


Figura 1.2: Opciones de gestión curricular en SICA

Este Trabajo de Fin de Máster pretende abordar esta carencia, permitiendo así que un investigador que tenga generado un CVN, con cualquier editor existente, pueda importar dichos datos a SICA.

Por lo tanto, el objetivo del proyecto será desarrollar una herramienta con la capacidad de importar elementos incluidos en el CVN a la plataforma SICA. Para ello se permitirá que un usuario introduzca un documento CVN en la herramienta, extrayendo todos los posibles datos curriculares que se incluyen. Tras esto, será posible seleccionar los elementos que desean ser importados a SICA, mostrando los campos acordes a la plataforma SICA para ese elemento y los valores obtenidos del CVN.

Para cumplir este objetivo será necesario realizar los siguientes pasos:

- **Extracción de información del CVN:** El documento CVN se presenta en formato PDF-XML, es decir, un fichero PDF con un fichero XML incrustado. Es en el archivo XML, conocido como CVN-XML, en donde se encuentra la información del currículum, regulada a partir de campos que siguen unas especificaciones técnicas. Esta información sigue una norma que establece las especificaciones que tienen seguir todos los CVN para que tengan cohesión en un único formato.
- **Análisis de elementos en SICA:** Cada elemento que puede ser añadido a SICA tiene unos campos específicos y que pueden no corresponderse con los que están incluidos en el CVN. Por lo tanto, es importante conocer todos los posibles campos de cada elemento y observar si su valor puede ser obtenido del CVN.
- **Análisis de tráfico en SICA:** Será necesario realizar un análisis mediante ingeniería inversa del tráfico HTTPS utilizado para añadir elementos en la plataforma SICA. Se estudiarán todos los datos que

se envían y se reciben en el proceso de añadir un nuevo elemento a la plataforma, estudiando las características del tráfico para cada uno.

- **Importación de datos a SICA:** Conociendo las características de subida de todos los posibles elementos en SICA y con la información de todos los datos del CVN que se introduzca en la herramienta, será posible realizar una importación de todos los datos del CVN a la plataforma.

Realizando estos pasos se pretende obtener la información necesaria para desarrollar una herramienta que permita importar a SICA cualquier elemento encontrado en un determinado CVN.

A continuación, se describe la estructura de esta memoria. El capítulo presente introduce las nociones básicas de este proyecto, presentando la carencia de SICA a la hora de importar el formato CVN y exponiendo la posible solución al desarrollar una herramienta que si lo permita. El capítulo 2 trata muestra el estado del arte, dando una perspectiva global del estado en el que se encuentra la materia desarrollada en este proyecto. El capítulo tres presenta los requisitos especificados para el trabajo. En el capítulo cuatro se expone una planificación y un presupuesto estimado según las horas dedicadas a este proyecto. En el capítulo cinco se explicará toda la información posible sobre el formato CVN, indicando sus especificaciones técnicas. El capítulo sexto es el que se muestra toda la parte referente al diseño del proyecto y la manera en la que se ha implementado. En el capítulo siete se desarrollará la fase de prueba que se ha realizado, mostrando los diferentes pasos que se han seguido para asegurar el correcto funcionamiento de la herramienta. Por último, el capítulo ocho incluye las principales conclusiones y líneas futuras.

Capítulo 2

Estado del arte.

En este capítulo se desarrollará el estado del arte, dando una perspectiva global del área con la que se trabaja en el proyecto.

Como es lógico, es más fácil gestionar un sistema científico si está bien estructurado e integrado. Los Curriculum Vitae de los investigadores son documentos que contienen datos individuales, instituciones y trabajos que se traducen en una gran fuente de información sobre la actividad científica de un país. Sin embargo, si esta información no es guardada apropiadamente no es posible utilizar este conocimiento para realizar diversos estudios. En lo que se refiere al uso del CV de los investigadores, el sistema Europeo de ciencia e investigación está caracterizado por dos características importantes: estandarización y el incremento de la digitalización. La estandarización de CV permite a las organizaciones mejorar sus procesos administrativos, así como la evaluación y gestión de subvenciones y financiación. Por otro lado, el incremento de la digitalización quiere decir que las instituciones públicas de investigación y las agencias de subvenciones utilizan nuevas tecnologías para obtener y guardar los CV de los investigadores. Un gran número de universidades y organizaciones públicas, como SICA, hacen uso de plataformas electrónicas que permiten al investigador introducir sus datos curriculares y permitir a las organizaciones usar esa información para diferentes objetivos de gestión y evaluación. Sin embargo, esta estandarización de los CV está aun bajo investigación y desarrollo, aunque cada vez se extiende más a lo largo del mundo debido a su gran utilidad.

En el caso de la plataforma SICA a la que se atañe este proyecto, la importación de un documento en formato CVN no es posible debido a su no implementación ni certificación por parte del FECYT. Sin embargo, existen diferentes plataformas que sí permiten la importación de un fichero CVN para registrar todos los datos.

Algunas entidades hacen uso del formato CVN, para su exportación, importación o ambas opciones. Accediendo a la dirección web <https://cvn.fecyt.es/servicios-cvn/instituciones-adaptadas-a-cvn>, es posible

observar el listado de instituciones que permiten incorporar un currículum en formato CVN directamente en la aplicación informática correspondiente. En esta página, también se encuentra a disposición el procedimiento de cada institución para generar el CVN desde su sistema.

En la siguiente figura se puede ver un extracto del listado proporcionado por FECYT, en el cual aparecen diferentes instituciones asociadas a su capacidad para exportar e importar CVN, incluyendo el número de versión del formato:

Entidad	Tipo de entidad	Sistema	Importación	Exportación
AGUA Y ESTRUCTURAS S.A.	Otros	SICA		v1.2.5
BASQUE CENTER FOR MACROMOLECULAR DESIGN & ENGINEERING, POLYMAT FUNDazioA	Otros	Fundanet		v1.4.0
BASQUE CENTRE FOR CLIMATE CHANGE	Otros	Fundanet		v1.4.0
CENTRO ANDALUZ DE BIOLOGÍA DEL DESARROLLO	Otros	SICA		v1.2.5
CENTRO ANDALUZ DE BIOLOGÍA MOLECULAR Y MEDICINA REGENERATIVA	Otros	SICA		v1.2.5
CENTRO ANDALUZ DE INNOVACIÓN Y TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES	Otros	SICA		v1.2.5
CENTRO DE EXPERIMENTACIÓN EL ARENOSILLO (INTA)	OPI	SICA		v1.2.5
CENTRO DE INNOVACIÓN ANDALUZ PARA LA CONSTRUCCIÓN SOSTENIBLE	Otros	SICA		v1.2.5
CENTRO DE INVESTIGACIÓN COOPERATIVA EN NANOCIENCIAS (CIC nanoGUNE)	Otros	Fundanet		v1.4.0
CENTRO DE INVESTIGACIÓN Y CALIDAD AGROALIMENTARIA DEL VALLE DE LOS PEDROCHES	Otros	SICA		v1.2.5
CENTRO DE REGULACIÓ GENÓMICA	Otros	Fundanet		v1.4.0
CENTRO TECNOLÓGICO AVANZADO DE LA PIEDRA	Otros	SICA		v1.2.5
ESCUELA ANDALUZA DE SALUD PÚBLICA	Otros	SICA		v1.2.5
ESCUELA DE ESTUDIOS ÁRABES (CSIC)	OPI	SICA		v1.2.5
ESCUELA DE ESTUDIOS HISPANO-AMERICANOS (CSIC)	OPI	SICA		v1.2.5
ESTACIÓN BIOLÓGICA DE DOÑANA (CSIC)	OPI	SICA		v1.2.5
ESTACIÓN EXPERIMENTAL DE ZONAS ÁRIDAS (CSIC)	OPI	SICA		v1.2.5
ESTACIÓN EXPERIMENTAL DEL ZAIDÍN (CSIC)	OPI	SICA		v1.2.5
ESTACIÓN EXPERIMENTAL LA MAYORA (CSIC)	OPI	SICA		v1.2.5
F.I.B. HOSPITAL INFANTIL UNIVERSITARIO NIÑO JESUS	Otros	iMarina		v1.3.0
FUNDACIÓ HOSPITAL UNIVERSITARI VALL D'HEBRON - INSTITUT DE RECERCA	Otros	CVNVHIR (iMarina)		v1.3.0
FUNDACIÓ INSTITUT D'INVESTIGACIÓ SANITÀRIA ILLES BALEARS	Otros	Fundanet		v1.4.0
FUNDACIÓ SANT JOAN DE DÉU	Otros	Fundanet		v1.4.0
FUNDACIÓN ANDALTEC	Otros	SICA		v1.2.5
FUNDACIÓN ANDALUZA BETURIA PARA LA INVESTIGACIÓN EN SALUD	Otros	SICA		v1.2.5

Figura 2.1: Extracto del listado de instituciones que permiten exportar e importar CVN

En ese listado aparecen 122 instituciones certificadas, de las cuales todas permiten la exportación al formato CVN, pero solo 15 de ellas permiten la importación. De ese listado, las instituciones que permiten la importación del formato CVN son las siguientes:

- Instituto de Investigación Sanitaria Gregorio Marañón
- Instituto de Investigación Sanitaria La Fe - Fundación para la Investigación del Hospital Universitario Politécnico La Fe
- Instituto de Salud Carlos III
- Junta de Extremadura
- Ministerio de Economía y Competitividad
- Universidad de Alcalá
- Universidad de Burgos

- Universidad de La Laguna
- Universidad de Las Palmas de Gran Canaria
- Universidad de León
- Universidad de Murcia
- Universidad de Navarra
- Universidad de Oviedo
- Universidad de Valladolid
- Universidad Rey Juan Carlos

Si la institución que se desea utilizar no se encuentra entre las certificadas, puede realizarse la importación del currículum mediante la herramienta del editor de FECYT. El editor CVN está disponible en <https://cvn.fecyt.es/editor/> y se trata de una aplicación on-line que pone a disposición de los investigadores la posibilidad de crear un currículum vitae desde cero o modificar el existente.

En la siguiente figura se muestra la pantalla de importación del editor de FECYT:

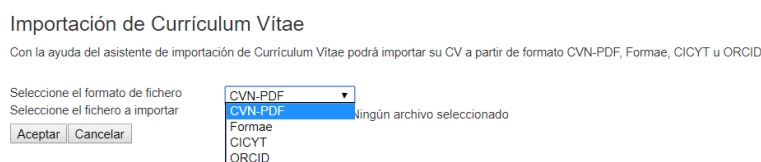


Figura 2.2: Importación de distintos formatos de CV a FECYT

Como se observa en la figura anterior, el editor de FECYT permite importar currículos completos en formato CVN-PDF, Formae, ORCID o CICYT.

Esta plataforma contiene, además, herramientas específicas para la importación de publicaciones desde diversos formatos, como el actual cargado en la aplicación, PubMed [5], WOK [6] y Scopus [7] y adaptar el CVN a la convocatoria interesada: crear uno reducido o ampliado. Es la herramienta recomendada en el caso de que una institución no permita todavía la conversión del currículum a CVN.

De hecho, numerosas plataformas dedicadas a la investigación e introducir reseñas curriculares como Web Of Science (WOS) o Scopus ya trabajan con el formato CVN, permitiendo exportar los datos a dicho formato. Sin embargo, aún no permiten la importación del formato CVN.

Existen otro tipo de herramientas que tienen como finalidad importar datos curriculares, por lo general, de forma manual. No existen demasiadas que permitan la importación de un CV, ya que este proceso requiere que el CV tenga un formato específico, como es el caso del CVN, y que sus campos estén adaptados a la plataforma a la que se requiere importar. Es el caso de Europass [8], un instrumento europeo que sirve para velar por la transparencia de las cualificaciones y las competencias, desarrollado en el año 2004, con objeto de ayudar a los ciudadanos a presentar sus capacidades y cualificaciones de una manera sencilla y fácilmente comprensible en Europa. El Europass está formado por cinco documentos, uno de ellos es el Currículum Vitae Europass. Este Currículum, parecido al Currículum Vitae Normalizado, es un modelo estandarizado común europeo de CV que posibilita a los individuos presentar de modo sistemático y cronológico sus calificaciones y competencias. Se trata por lo tanto de un CV con un formato propio, que además puede ser importado en la plataforma de Europass y que es promocionado por SEPIE (Servicio Español Para la Internacionalización de la Educación).

En la figure 2.3, obtenida de la propia plataforma de Europass se pueden ver las distintas opciones para trabajar con CV:

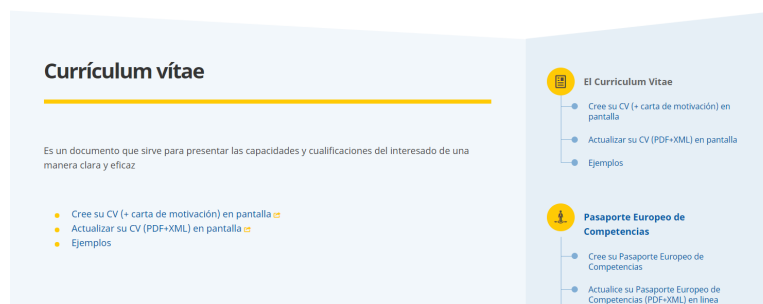


Figura 2.3: Plataforma Europass

Otro conocido estándar utilizado para el intercambio de información científica es el esquema METS [9], utilizado para la codificación electrónica de un documento u objeto dentro de una biblioteca digital. Se utiliza para ello un esquema de lenguaje XML de la World Wide Web Consortium (W3C) [10]. La norma es mantenida como parte de los estándares MARC [11] de la Biblioteca del Congreso de Estados Unidos y está siendo desarrollada como una iniciativa de la Digital Library Federation (DLF) [12]. Con este estándar pueden identificarse contenidos de un objeto digital y expresar la estructura de este contenido, además de coordinar el comportamiento de las distintas partes que constituyen el documento.

Por otro lado, existen otro tipo de formatos de plataformas y formatos de CV ajenos a España y la UE, como es el caso de CVar del Ministerio de Ciencia, Tecnología e Innovación Productiva de Argentina [13]. El CVar es

el Registro Unificado y Normalizado a nivel nacional de los Datos Curriculares del personal científico y tecnológico que se desempeña en las distintas instituciones argentinas. Esta plataforma permite la importación de datos de otros formatos de CV, como es el caso de SIGEVA (Sistema Integral de Gestión y Evaluación) [14].

Como se ha podido comprobar, el número de plataformas que permiten la importación de un CV son realmente escasas, en especial aquellas que permiten la importación de un CVN, siendo solo 15 plataformas las que poseen esta funcionalidad. Existen plataformas dedicadas a introducir reseñas curriculares o a publicar investigación que permiten otro tipo de importación, como es el caso de ORCID [15], una plataforma que proporciona un código alfanumérico para identificar de manera única a científicos y otros autores académicos. ORCID permite importar trabajos a su plataforma a través de enlaces o búsquedas en otros sistemas en los cuales ya existan esos trabajos. Por ejemplo, permite importar trabajos de sitios como Scopus, ResarcherID, PubMed o Airiti. La lista completa de sitios disponibles para la importación de trabajos en ORCID se puede encontrar en <https://support.orcid.org/knowledgebase/articles/188278-add-works-by-direct-import-from-other-systems>.

En la figura 2.4, obtenida de la propia página de ORCID, se observa cómo es posible dicha importación.

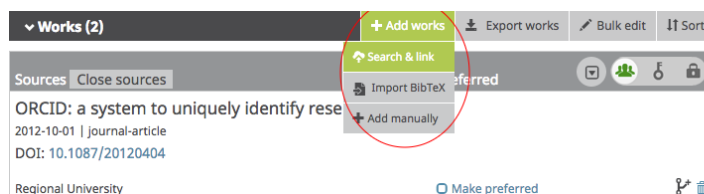


Figura 2.4: Importación de trabajos en ORCID

ResearcherID [16], un sistema de identificación de autores científicos, permite la importación del formato RIS (Research Information Systems). Este formato es utilizado en numerosas librerías digitales como IEEE Xplore, Scopus o ACM, que permiten la exportación de citas a RIS.

Por otro lado, la plataforma EndNote, conocida por ser un paquete informático de gestión de referencias, permite la importación de citas o listados bibliográficos en diferentes formatos, como es el caso de RIS, Google Books, PubMed o una propia importación del formato EndNote.

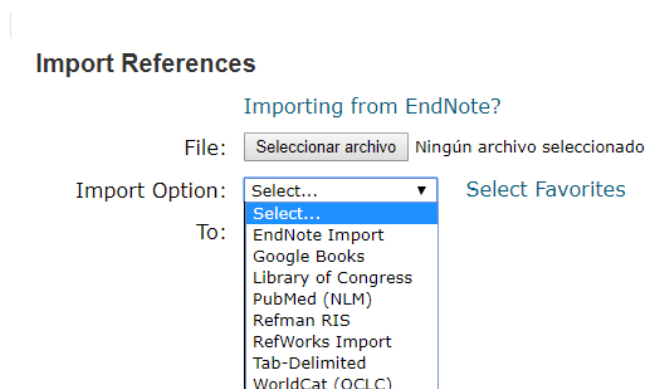


Figura 2.5: Importación en diferentes formatos a EndNote

Como se ha podido observar en el estado del arte de este proyecto, la mayoría de plataformas dedicadas a la investigación y reseñas curriculares no permiten importación de un Currículum Vitae, sino que están diseñadas para poder importar citas o referencias bibliográficas. Es por eso, que es necesario un formato de normalización de CV como es el caso de CVN para implementar en diferentes plataformas la funcionalidad de importar los datos almacenados en un CV.

Capítulo 3

Requisitos.

En este capítulo se pretende llevar a cabo un análisis del proceso de diseño realizado para este proyecto, incluyendo, de la manera más completa posible, los requisitos relacionados con la funcionalidad y las decisiones previas que se deben cumplir. Estos requisitos son importantes, ya que proporciona documentación sobre el contenido del proyecto durante la fase de diseño.

A continuación se listarán los requisitos mínimos que serán necesarios para el funcionamiento del proyecto, y que serán útiles a la hora de tener presente las ideas del diseño:

- **Obtención del CVN-XML incrustado en el CVN-PDF:**

Será necesario que la herramienta realizada sea capaz de obtener el XML incrustado en el CVN-PDF. En dicho XML se encuentra toda la información del currículum que se desea importar, por lo que es un requisito fundamental del proyecto.

- **Análisis de la información del CVN-XML:**

La herramienta deberá poder analizar los distintos campos de los que se compone el currículum en el formato CVN y mostrar por pantalla los que sean compatibles con los elementos de CV incluidos en la plataforma SICA. Para ello, se obtendrán del CVN aquellos valores de campos que sean compatibles con los que incluye SICA para un determinado elemento, ya que existen ciertos campos del CVN que pueden no estar presentes en la plataforma.

- **Análisis de campos y tráfico HTTPS de SICA:**

Será totalmente necesario conocer todos los posibles campos para los posibles elementos que pueden ser añadidos a SICA, observando qué valores de los mismos pueden ser obtenidos del CVN. Adicionalmente,

será necesario analizar el tráfico a la hora de añadir de un elemento en la plataforma SICA. Este proceso deberá incluir el análisis del tráfico en el acceso a la plataforma con las credenciales de usuario y el tráfico que se genera tras añadir un nuevo elemento, observando todos los datos que se envían y se reciben.

- **Importación de CVN-PDF o CVN-XML a SICA:**

El requisito más importante del proyecto, y en el que está basado, será la de realizar una herramienta con la capacidad de importar un documento CVN (en formato PDF o XML) a la plataforma SICA, haciendo uso para ello de los requisitos nombrados anteriormente.

- **Utilidad y facilidad**

Uno de los factores más importantes en la herramienta a desarrollar debe ser su utilidad para los distintos investigadores que la utilicen, por lo que es importante que se consiga la importación a SICA de la manera más completa posible, intentado que el usuario realice el menor número de acciones posibles.

- **Interfaz sencilla e intuitiva**

El objetivo de la herramienta es la importación de ficheros en formato CVN a SICA, siendo el principal objetivo su utilidad como se comentaba anteriormente. Por lo tanto, la interfaz de la herramienta debe ser sencilla e intuitiva de modo que no sea complicado utilizarla por los usuarios que lo deseen.

- **Herramienta tolerante a fallos**

La herramienta desarrollada debe ser tolerantes a fallos e informar al usuario de posibles errores en la importación de ficheros CVN a SICA, indicando los elementos que no pudieron ser importados.

Además de estos requisitos, una decisión previa al diseño ha sido desarrollar la herramienta a través de una interfaz web, por lo que será necesario utilizar lenguaje HTML, CSS y JavaScript para el lado del cliente, es decir, lo que visualizará el usuario. Para la parte del servidor en donde se realizará todo el proceso de análisis del CVN y la importación a SICA se utilizará el lenguaje PHP. Esta decisión se debe a que una herramienta desarrollada sobre una interfaz web ofrece un fácil acceso para los usuarios que deseen importar su CVN, en lugar de instalar la herramienta en un sistema operativo.

Capítulo 4

Planificación y presupuesto.

En el presente capítulo se tratarán los aspectos relacionados con la planificación y el presupuesto del proyecto, haciendo una estimación de costes.

En primer lugar, se tratarán las diferentes etapas de las que ha constado el proyecto y el número de horas dedicado a cada etapa. Para ello, se realizará un diagrama de Gantt que representará gráficamente la evolución del proyecto. Además, posteriormente se identificarán los recursos utilizados, clasificándolos en recursos humanos, hardware y software.

Por otro lado, se realizará una aproximación del presupuesto total necesario para llevar a cabo el proyecto, basándose en los recursos utilizados y el estudio de costes realizado.

4.1. Planificación

En esta sección se describirá la planificación que se ha adoptado para la realización del proyecto.

Tarea 1: Revisión bibliográfica

En esta primera etapa se realizará una investigación sobre qué es el formato CVN y la forma de utilizarlo en diferentes plataformas, además de documentación oficial proporcionada por FECYT que revele información útil sobre la norma CVN. Además, se investigarán todas las plataformas que permitan la importación del formato, concluyendo la mejor manera de realizar este proyecto basado en las posibles soluciones que se encuentren.

Tarea 2: Definición de requisitos y elección de plataformas

Esta segunda etapa consistirá en planificar que requisitos serán los necesarios según el problema planteado, y cuál podría ser la plataforma adecuada en la que desarrollar la herramienta que permita importar ficheros CVN a SICA.

Tarea 3: Aprendizaje sobre el formato CVN

En esta etapa se realizará una investigación sobre el formato CVN, tanto su modalidad CVN-PDF como CVN-XML. Se determinará como obtener el fichero XML de un CVN-PDF y se intentarán analizar todos los campos y códigos posibles que pueden incluirse en un fichero CVN-XML.

Tarea 4: Aprendizaje sobre la plataforma SICA

En esta etapa se profundizará en la plataforma SICA, aprendiendo su funcionamiento, los posibles elementos que pueden ser añadidos y la manera de hacerlo.

Tarea 5: Diseño del proyecto

La etapa de diseño es de gran importancia, ya que consistirá en planificar como desarrollar la arquitectura final del proyecto, según la solución adoptada al problema planteado.

Tarea 6: Análisis de todos los posibles campos del formato CVN

En esta tarea se realizará un análisis de todos los posibles campos que pueden ser incluidos en el formato CVN. Para ello, se generarán a través del editor de FECYT todos los posibles elementos, con la mayor cantidad de información posible, y se analizará el fichero generado. Se determinarán todos los posibles campos relacionados con cada elemento generado, para posteriormente observar cuales son válidos para ser utilizados en la plataforma SICA.

Tarea 7: Análisis de los elementos y campos en SICA

Esta tarea requiere conocer todos los elementos que pueden ser generados en SICA y exportarlos al formato CVN. Una vez realizada la exportación, se analizará el CVN generado, identificando a que elemento del CVN se corresponde cada uno de los exportados en SICA. Además, será necesario observar todos los posibles campos de cada elemento en SICA, observando que valores pueden ser obtenidos a partir de un CVN. Esta tarea es la más tediosa, pues requiere realizar subidas de elementos uno por uno y examinar el CVN generado posteriormente.

Tarea 8: Análisis del tráfico en SICA

Esta tarea está estrechamente relacionada con la tarea anterior, ya que en el proceso de generar cada elemento en SICA deberá estudiarse el tráfico que se genera en el proceso, observando todos los datos que se envían en los formularios entre las distintas páginas.

Tarea 9: Creación de herramienta de importación de CVN a SICA

La herramienta para importación de CVN a SICA se realizará mediante programación en lenguaje PHP para la parte servidora y JavaScript, HTML y CSS para la parte visible por el cliente. Esta es la parte práctica del proyecto y el objetivo principal del mismo.

Tras conocer el formato de CVN y el estudio del análisis del tráfico en SICA, se realizará una herramienta con la capacidad de analizar un CVN-PDF o CVN-XML y analizar todos los elementos que contiene, con toda la información posible. Una vez analizado en su totalidad, se mostrarán por pantalla al usuario todos los elementos que se han encontrado y los valores para cada campo, dando la opción de elegir aquellos elementos de su CV que desea importar a la plataforma SICA.

Tarea 10: Fase de pruebas

En esta etapa se realizarán las pruebas necesarias para las implementaciones realizadas en el proyecto. Se tendrán que corregir los fallos detectados en las etapas de implementación y por otro lado, se comprobará que el sistema cumple con su propósito correctamente.

Tarea 11: Memoria técnica del proyecto

En esta etapa final se realizará una memoria técnica sobre el presente proyecto, en el que se explicarán todos los pasos detalladamente, documentando de la mejor manera posible todos los aspectos que puedan tener relevancia para este proyecto. Con esta memoria se pretende informar al lector de cualquier detalle que pueda resultar de interés sobre la elaboración del proyecto y la información referente al mismo.

Tarea 12: Exposición del Trabajo de Fin de Máster

Como última etapa del proyecto, se realizará una exposición frente a un tribunal defendiendo el presente trabajo y sus características.

Una vez que han quedado identificadas las tareas realizadas en este proyecto, se hará una planificación dependiendo del tiempo estimado que tomaría realizar cada una de las tareas. Esta distribución se puede ver mediante un diagrama de Gantt (Ver figura 4.1), donde se hace una estimación de los días planificados a cada una de las tareas.

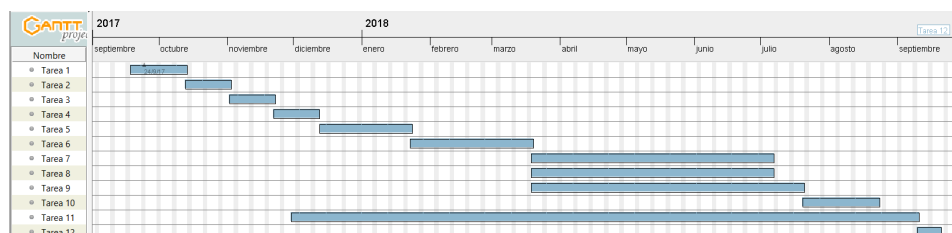


Figura 4.1: Diagrama de Gantt

Como se puede observar en el diagrama de Gantt, existen tareas que se realizan paralelamente a otras. Es el caso del análisis de los elementos, campos y tráfico en SICA (tareas 7 y 8) y la del desarrollo de la herramienta (tarea 9) ya que están estrictamente relacionadas. Además, en todo el proceso de desarrollo de este proyecto se realizará conjuntamente la memoria del mismo, junto con el resto de tareas, documentando la mayor cantidad de información posible.

4.2. Presupuesto

En esta sección se tratará el tema del presupuesto del proyecto, en la que se listarán los recursos utilizados y el precio estimado correspondiente a cada uno.

4.2.1. Recursos humanos

Los recursos humanos son aquellos en los que ha sido necesaria la intervención humana para la realización del proyecto. En el caso del presente trabajo los recursos humanos son:

- Juan Carlos Angulo Santos: Autor del proyecto, graduado en Ingeniería de Telecomunicaciones y alumno del Máster en Ingeniería de Telecomunicaciones.
- Jorge Navarro Ortiz: Tutor del proyecto y profesor contratado doctor de la Universidad de Granada del Departamento de Teoría de la Señal, Telemática y Comunicaciones.

Teniendo en cuenta que el precio de un ingeniero graduado en ingeniería de tecnologías de telecomunicación es de aproximadamente 25 euros/hora y el de un doctor de la Universidad de Granada es de aproximadamente 50 euros/hora, podemos hacer un cálculo estimado de los recursos humanos invertidos en el proyecto. Aproximadamente, el tutor del proyecto, entre tutorías y revisión del trabajo, habrá invertido unas 15 horas.

En la siguiente tabla (véase tabla 4.1) podemos observar el número de horas dedicado a cada tarea acorde a tal y como se vio en el diagrama de

Gantt. En otra columna se mostrará el precio estimado para dicha tarea y finalmente se concluirá con el precio total que tomaría realizar este proyecto únicamente haciendo uso de los recursos humanos.

Tarea	Tiempo (Horas)	Coste (Euros)
Revisión bibliográfica	20	500
Definición de requisitos y elección de herramientas/plataformas	20	500
Aprendizaje sobre el formato CVN	30	750
Aprendizaje sobre la plataforma SICA	30	750
Diseño del proyecto	40	1000
Análisis de todos los posibles campos del formato CVN	50	1250
Análisis de los elementos y campos en SICA	70	1750
Análisis del tráfico en SICA	70	1750
Creación de herramienta de importación de CVN a SICA	90	2250
Fase de pruebas	30	750
Memoria técnica del proyecto	50	1250
Exposición del TFM	25	625
Trabajo de ingeniero doctor	15	375
Total	540	13500

Tabla 4.1: Estudio de costes para los recursos humanos.

Como se observa, según la estimación realizada, el presupuesto total para los recursos humanos será de 13500 euros.

4.2.2. Recursos hardware

Respecto a los recursos hardware utilizados para este proyecto, únicamente ha sido necesario un Ordenador portátil ASUS R510, con procesador Intel Core i7 6700 HD y 12 GB de memoria RAM, con un coste de 600 euros. Es importante un buen equipamiento informático que permita la ejecución del proyecto en un tiempo apropiado, aumentando la productividad a la hora de trabajar.

Considerando que la vida útil de un portátil es de cinco años, y dado que este proyecto ha tenido una duración aproximada de un año, el presupuesto equivalente del portátil utilizado será de aproximadamente 120 euros.

4.2.3. Recursos software

Estos son los recursos software que se han utilizado:

- Sublime Text: Es un editor de texto gratuito dedicado a código, utilizado en este proyecto para la programación en PHP. Además, se ha utilizado para visualizar los archivos CVN-XML obtenidos del CVN-PDF.
- XAMPP: Es un servidor independiente de plataforma de software libre que se ha utilizado para emular el servidor en una página web y hacer distintas pruebas.
- TeXstudio: Es un editor de LaTeX de código abierto que se ha utilizado para redactar la memoria del proyecto.
- GanttProject: Es el programa utilizado para realizar el diagrama de Gantt que se mostraba anteriormente.
- SourceTree: Es un cliente GUI utilizado para manejar repositorios GIT que se ha empleado para realizar copias de seguridad del proyecto. Además, permite de forma sencilla observar los cambios producidos en cada versión que se ha subido.
- Mitmproxy: Se trata de un proxy *Man in the middle*, utilizado para examinar el tráfico HTTPs al generar elementos en SICA. Se ha utilizado una versión para web conocida como mitmweb.

En el caso de los recursos software, se ha buscado que todos los programas utilizados tuviesen un coste gratuito, con el propósito de hacer este proyecto lo más económico posible.

4.2.4. Presupuesto final

El presupuesto final será la suma del estudio de costes realizado para los recursos humanos, hardware y software. Puede verse el precio estimado para cada sección en la tabla 4.2:

Recursos	Presupuesto (Euros)
Recursos humanos	13500
Recursos hardware	120
Recursos software	0
Total	13620

Tabla 4.2: Presupuesto final del proyecto.

En definitiva, el presupuesto total del proyecto será de 13620 euros, considerando todos los aspectos anteriores.

Capítulo 5

Currículum Vitae Normalizado.

En este capítulo se detallará qué es el estándar CVN, analizando su procedencia, usos y ventajas. Además, se describirá la norma CVN que incluye las especificaciones técnicas del formato.

5.1. El estándar CVN

El Currículum Vitae normalizado, conocido por sus siglas CVN, es un proyecto de la Federación Española para la Ciencia y la Tecnología (FECYT) y del Ministerio de Economía, Industria y Competitividad del gobierno español, para la implantación de las tecnologías de la información en la gestión de la actividad investigadora, con el objetivo de evitar que el investigador tenga que completar de nuevo su CV en los formularios o bases de datos de las convocatorias. Este proyecto comenzó a implantarse en 2006 en universidades y organismos de investigación, adaptando la información ya cargada de los investigadores al estándar CVN.

El CVN es una Norma estándar que define un mismo formato de presentación de los datos curriculares de los investigadores y que posibilita la interoperabilidad con las diferentes bases de datos de las instituciones. Tiene como objetivo que los investigadores puedan presentar su currículum en un formato digital único y común que les permite actualizar de forma automática su información curricular, sin tener que volver a introducirla de nuevo. Por lo tanto, CVN facilita la presentación del CV de una forma unificada en las convocatorias de ayudas públicas del Plan Estatal y autonómicas.

CVN es utilizado por investigadores, instituciones con personal investigador y gestores públicos de investigación. De hecho, es obligatorio en las convocatorias de recursos humanos para investigación que se gestionan desde la Secretaría de Estado I+D+i (MINECO) y en las ayudas de recursos humanos, redes y proyectos de la Acción Estratégica en Salud (AES).

Además, es recomendado en las convocatorias del Plan Estatal de I+D+i por el Ministerio de Economía y Competitividad.

El formato CVN cuenta con los siguientes elementos, que sirven como herramientas de soporte y apoyo del mismo:

- **Una Norma:** Se trata de un conjunto de especificaciones técnicas que regulan los campos del currículo estándar.
- **Una página web:** La página web <https://cvn.fecyt.es/> contiene información del proyecto, así como de las instituciones que se van sumando al mismo y un acceso privado para los administradores de las instituciones que se adaptan al formato CVN.
- **Un editor de currículos CVN:** FECYT pone a disposición de los investigadores una herramienta *online* (<https://cvn.fecyt.es/editor/>) que da la posibilidad de crear un currículum desde cero o modificar uno existente. Además, esta herramienta contiene elementos específicos para la importación de CVs en otros formatos o para la importación de publicaciones desde diferentes plataformas (Scopus, PubMed...).
- **Servicio web:** Este servicio facilita que todos los CVN originados dispongan de un sello de calidad de FECYT mediante una firma electrónica. Este sello de calidad implica que los CVN se han generado conforme a la Norma.
- **Modo de lectura de CVN:** Este módulo permite convertir un fichero en formato XML con lenguaje técnico en un PDF que sea legible y presentable.
- **Soporte y certificación de las instituciones:** Servicio para adaptar sistemas de gestión curricular a la Norma CVN.

Desde el punto de vista tecnológico, un currículum vitae en formato CVN es un fichero electrónico en PDF (CVN-PDF) con tecnología CVN-XML incrustada, de tal manera que la información normalizada que contiene puede ser tratada electrónicamente por distintas bases de datos; con la intención de evitar que los investigadores tengan que teclear su currículo en los distintos formularios o aplicaciones que se utilizan para la gestión curricular y en las diferentes convocatorias de ayudas.

Este fichero en formato CVN-PDF puede ser archivado, visualizado o impreso por el investigador, pero a efectos de seguridad e integridad de la información, cualquier modificación o manipulación posterior a su obtención desde una aplicación informática lo invalidará. Por lo tanto, cualquier modificación o subsanación de error que se desee, deberá realizarse previamente en la aplicación curricular de una institución o a través del editor CVN de FECYT.

Las ventajas del uso del servicio CVN aplica a investigadores, instituciones, gestores de convocatorias de ayudas y, por ende, a los evaluadores de dichas ayudas, ya que:

- Reduce la carga de trabajo que necesita un investigador para mantener actualizado su CV, manteniendo unificado el formato de los CVs generados por diferentes instituciones públicas.
- Mejora las tareas administrativas de la gestión curricular de los solicitantes a las diferentes convocatorias públicas. Facilita y reduce el tiempo de los procesos de evaluación y acreditación del personal investigador, permitiendo a todo el sistema emplear un mismo conjunto de especificaciones para el intercambio de información curricular.
- Simplifica la tramitación electrónica y contribuye al desarrollo de la e-administración, lo que redundará en una disminución de los trámites presenciales.

Por otro lado, FECYT pone a disposición de las instituciones un manual de certificaciones [17] en el cual se muestra el proceso para la certificación de instituciones que quieran adaptarse a CVN. Se trata del proceso de validación informática que, por medio de unas pruebas, permite a FECYT dar las garantías suficientes de la generación o lectura de currículos en el formato CVN. Este proceso conlleva un desarrollo tecnológico por parte de las instituciones y se realiza en dos fases: exportación e importación.

En este manual se explican los pasos a realizar para la certificación a CVN de una institución. El primero de los pasos es contactar con FECYT mediante correo electrónico y concertar así una primera reunión, en donde explica y detalla todo el proceso tanto desde el punto de vista funcional como tecnológico del proyecto. Seguidamente, FECYT facilitará el acceso a la institución a la parte privada de la web CVN, entre otros. Al autenticarse y acceder al área privada la institución dispondrá de toda la documentación relativa a la norma CVN y los pasos a seguir para la homologación de instituciones. Esto quiere decir que la información sobre el formato de la tecnología CVN-XML es privada y pertenece a FECYT, por lo que no existe información pública sobre el formato.

5.2. La norma CVN

La norma CVN, como se comentó en la sección anterior, es el conjunto de especificaciones técnicas que regulan todos los campos del currículum estándar, sin tener ninguno de ellos un carácter obligatorio. Por lo tanto, es el reglamento que deben seguir todos los currículos adaptados al formato CVN para que tengan cohesión en un único formato. Esta Norma comprende

la información que un investigador, tecnólogo o innovador puede necesitar para reseñar su trayectoria e incluye su interacción con otras dedicaciones.

La Norma CVN pretende sentar las bases para la creación de una red de intereses curriculares comunes para el Sistema Español de Ciencia Tecnología e Innovación. Con este objetivo, la FECYT ha elaborado, en colaboración con otras instituciones participantes y consultadas, un documento normativo para la estandarización de los Currículos Vítae de I+D+i.

La implantación de este documento es un proceso largo, que conlleva un método específico. Su aplicación debe estar planificada y documentada fomentando su evolución y mejora. FECYT asegura, mediante la certificación, que los currículos de las instituciones adaptadas a CVN cumplen con la especificación de esta Norma.

En esta sección se detallará como está realizada esta Norma CVN, que se encuentra implementada en formato XML, incrustado en un CVN-PDF como se ha comentado en diversas ocasiones. En la figura 5.1 se observa un ejemplo de un CVN-XML en donde se muestra el formato que sigue:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:CVN xmlns:ns2="http://cv.normalizado.org/cvn">
  <Version>
    <VersionID>
      <Date obligatory="true" code="000.020.000.020">
        <Item>2018-02-11+01:00</Item>
      </Date>
      <CVLanguage>
        <LanguageName obligatory="false" code="000.020.000.070">
          <Item>Español</Item>
        </LanguageName>
        <LanguageCode obligatory="false" code="000.020.000.070">
          <Item>spa</Item>
        </LanguageCode>
      </CVLanguage>
    </VersionID>
  </Version>
  <Agent>
    <Identification>
      <PersonalIdentification>
        <GivenName code="000.010.000.020">
          <Item>Juan Carlos</Item>
        </GivenName>
        <FirstFamilyName code="000.010.000.010">
          <Item>Angulo</Item>
        </FirstFamilyName>
        <SecondFamilyName code="000.010.000.010">
          <Item>Santos</Item>
        </SecondFamilyName>
      </PersonalIdentification>
    </Identification>
  </Agent>
</ns2:CVN>
```

Figura 5.1: Ejemplo de formato CVN-XML

CVN-XML es un modelo abstracto con la capacidad de describir cualquier tipo de actividad profesional, actuando como un traductor universal y un contenedor de diferentes CVs. Un conjunto de lenguajes XSD [18] y va-

lores XML proporcionan toda la semántica y síntesis necesaria para aplicar el modelo CVN-XML.

Como se observa en la figura 5.1 el formato CVN se compone de diversos campos, que se estructuran de una forma ordenada. Cada *item* o elemento curricular que puede encontrarse en el CVN tiene unos atributos utilizados para la descripción de cada campo.

A continuación, se nombrarán los campos más importantes en el CVN-XML que permiten organizar la información curricular.

En primer lugar, el CVN-XML se define con un nodo principal:

```
<ns2:CVN xmlns:ns2="http://cv.normalizado.org/cvn">
```

Lo que engloba ese nodo incluye todos los campos del CVN y son los que se analizarán y validarán por cualquier plataforma que permita la importación del formato. Este nodo incluye otros tres nodos principales: *Version*, *Agent* y *CvnItem*. En el siguiente listado se define el uso de cada uno:

- **Version:** Incluye información sobre la identificación del currículum.
- **Agent:** Identifica a la persona que ha generado el currículum, conteniendo datos personales como nombres, apellidos, DNI o fecha de nacimiento.
- **CvnItem:** Este nodo hace referencia a un elemento que forma parte del currículum. Puede repetirse tantas veces como sea necesario, uno por cada reseña curricular que contenga el CVN.

En la figura 5.2 se observa un ejemplo de los nodos principales que contendría un CVN-XML, en el cual se muestra el nodo *Version*, *Agent* y tres nodos *CvnItem* ya que únicamente se han introducido tres reseñas curriculares para ese CVN. El contenido de cada nodo ha sido comprimido con el objetivo de que se visualice correctamente el formato de un CVN típico.



```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:CVN xmlns:ns2="http://cv.normalizado.org/cvn">
  <Version> ...
</Version>
  <Agent> ...
</Agent>
  <CvnItem> ...
</CvnItem>
  <CvnItem> ...
</CvnItem>
  <CvnItem> ...
</CvnItem>
</ns2:CVN>
```

Figura 5.2: Nodos principales del CVN-XML

Cada campo del formato CVN se identifica, además de con el nombre del nodo, con un código único para cada uno. Este código se compone de doce números, divididos en cuatro subgrupos separados por un punto, con formato *XXX.XXX.XXX.XXX*.

A continuación, se muestran todos los campos que pueden contener los nodos *Version* y *Agent*, y algunas de las que podría contener el nodo *CvnItem*, ya que los campos en este último varían según el tipo de elemento curricular que contiene el CVN.

Es importante destacar que este análisis de nodos y códigos ha sido realizado mediante la creación de numerosos CVNs a través del editor de FECYT y sin una documentación oficial, por lo que es probable que uno de estos nodos contenga otros que no ha sido posible identificar.

Campos del nodo *Version*

El nodo *Version* se corresponde con el grupo de códigos 000.020.000.000, que a su vez contiene otro llamado *VersionID*. El nodo *VersionID* puede contener los siguientes otros nodos:

- **VersionID**
 - **CVN_Identification** (000.020.000.010): Código de identificación de CVN
 - **Date** (000.020.000.020): Fecha del documento
 - **CVLanguage**: Información de idioma
 - **LanguageName** (000.020.000.070): Nombre del idioma
 - **LanguageCode** (000.020.000.070): Código de idioma
 - **CodificationVersion** (000.020.000.080): Versión de codificación

- **CVUser:** Información de usuario de CV
 - **User** (000.020.000.030): Usuario de CV
 - **Password** (000.020.000.040): Contraseña de CV
 - **UpdateDate** (000.020.000.050): Fecha de actualización

Como se observa, todos los campos incluidos en el nodo *Version* comienzan con un código 000.020, ya que es el prefijo para este tipo de nodo.

En la figure 5.3 se puede observar un ejemplo del nodo *Version* incluido en el CVN-XML:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:CVN xmlns:ns2="http://cv.normalizado.org/cvn">
  <Version>
    <VersionID>
      <CVN_Identification code="000.020.000.010">
        <Item>843524660002800010000000000499</Item>
      </CVN_Identification>
      <Date code="000.020.000.020">
        <Item>2018-06-28</Item>
      </Date>
      <CVLanguage>
        <LanguageName code="000.020.000.070">
          <Item>Español</Item>
        </LanguageName>
        <LanguageCode code="000.020.000.070">
          <Item>spa</Item>
        </LanguageCode>
      </CVLanguage>
      <CodificationVersion code="000.020.000.080">
        <Item>1.3.0</Item>
      </CodificationVersion>
      <CVUser>
        <User obligatory="false" code="000.020.000.030" multiplicity="false">
          <Item>No_se_facilita_esta_información</Item>
        </User>
        <Password obligatory="false" code="000.020.000.040" multiplicity="false">
          <Item>No_se_facilita_esta_información</Item>
        </Password>
        <UpdateDate>
          <DayMonthYear obligatory="false" code="000.020.000.050" multiplicity="false">
            <Item>2018-06-15</Item>
          </DayMonthYear>
        </UpdateDate>
      </CVUser>
    </VersionID>
  </Version>
</ns2:CVN>
```

Figura 5.3: Campos para el nodo *Version*

Campos del nodo *Agent*

El nodo *Agent* se corresponde con el grupo de códigos 000.010.000.000. Este nodo contiene otros tres nodos importantes que pueden ser considerados como diferentes categorías. Estos son los nodos *Identification*, *Address* y *Contact*. En la siguiente lista se muestran todos los nodos posibles para el campo *Agent*:

■ Identification

- **PersonalIdentification**: Identificación del propietario del CVN
 - **GivenName** (000.010.000.020): Nombre del propietario
 - **FirstFamilyName** (000.010.000.010): Primer apellido del propietario
 - **SecondFamilyName** (000.010.000.010): Segundo apellido del propietario
 - **OfficialId**: Tipo de identificación y número
 - ◇ **DNI** (000.010.000.100): DNI del propietario
 - ◇ **NIE** (000.010.000.110): NIE del propietario
 - ◇ **Pasaport** (000.010.000.120): Pasaporte del propietario
 - **Nacionality** (000.010.000.040): Código que identifica la nacionalidad del propietario
 - **BirthDate** (000.010.000.050): Fecha de nacimiento
 - **BirthCountry** (000.010.000.060): Código que identifica el país de nacimiento
 - **BirthRegion** (000.010.000.070): Código que identifica la región de nacimiento
 - **BirthCity** (000.010.000.090): Código que identifica la ciudad de nacimiento
 - **Gender** (000.010.000.030): Código que identifica el género del propietario

■ Address: Dirección del propietario del CVN

- **City** (000.010.000.170): Ciudad
- **Streets** (000.010.000.140): Calle
- **OtherInformation** (000.010.000.150): Otra información
- **PostalCode** (000.010.000.160): Código postal
- **Region** (000.010.000.190): Región
- **CountryCode** (000.010.000.180): Código de país

■ Contact: Información de contacto del propietario del CVN

- **Telephone** (000.010.000.210): Teléfono fijo
- **Telephone** (000.010.000.240): Teléfono móvil
- **Fax** (000.010.000.220): Fax
- **InternetEmailAddress** (000.010.000.230): Correo electrónico
- **PersonalWeb** (000.010.000.250): Página web personal

En las figuras 5.4 y 5.5 se muestra un ejemplo real de cómo se visualiza el nodo *Agent* y todos sus campos en un CVN-XML:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:CVN xmlns:ns2="http://cv.normalizado.org/cvn">
  <Agent>
    <Identification>
      <PersonalIdentification>
        <GivenName obligatory="true" code="000.010.000.020" multiplicity="false">
          <Item>Juan Carlos</Item>
        </GivenName>
        <FirstFamilyName obligatory="true" code="000.010.000.010" multiplicity="false">
          <Item>Angulo</Item>
        </FirstFamilyName>
        <SecondFamilyName obligatory="false" code="000.010.000.010" multiplicity="false">
          <Item>Santos</Item>
        </SecondFamilyName>
        <OfficialId>
          <DNI obligatory="false" code="000.010.000.100" multiplicity="false">
            <Item>75679843P</Item>
          </DNI>
        </OfficialId>
        <Nationality obligatory="false" code="000.010.000.040" multiplicity="false">
          <Item>724</Item>
        </Nationality>
        <BirthDate obligatory="true" code="000.010.000.050" multiplicity="false">
          <Item>1993-08-03</Item>
        </BirthDate>
        <BirthCountry obligatory="false" code="000.010.000.060" multiplicity="false">
          <Item>724</Item>
        </BirthCountry>
        <BirthRegion>
          <Code obligatory="false" code="000.010.000.070" multiplicity="false">
            <Item>ES13</Item>
          </Code>
          <Name obligatory="false" code="000.010.000.070" multiplicity="false">
            <Item>Cantabria</Item>
          </Name>
        </BirthRegion>
        <BirthCity obligatory="false" code="000.010.000.090" multiplicity="false">
          <Item>Santander</Item>
        </BirthCity>
        <Gender obligatory="true" code="000.010.000.030" multiplicity="false">
          <Item>000</Item>
        </Gender>
      </PersonalIdentification>
    </Identification>
  </Agent>
</CVN>
```

Figura 5.4: Campos para el nodo *Agent* (I)

```

<Address>
  <City obligatory="false" code="000.010.000.170" multiplicity="false">
    <Item>Santander</Item>
  </City>
  <Streets obligatory="false" code="000.010.000.140" multiplicity="false">
    <Item>Avda Los Castros 40</Item>
  </Streets>
  <OtherInformation obligatory="false" code="000.010.000.150" multiplicity="false">
    <Item>Escuela de Telecomunicaciones</Item>
  </OtherInformation>
  <PostalCode obligatory="false" code="000.010.000.160" multiplicity="false">
    <Item>39005</Item>
  </PostalCode>
  <Region>
    <Code obligatory="false" code="000.010.000.190" multiplicity="false">
      <Item>ES13</Item>
    </Code>
    <Name obligatory="false" code="000.010.000.190" multiplicity="false">
      <Item>Cantabria</Item>
    </Name>
  </Region>
  <CountryCode obligatory="false" code="000.010.000.180" multiplicity="false">
    <Item>724</Item>
  </CountryCode>
</Address>
<Contact>
  <Telephone>
    <Number obligatory="true" code="000.010.000.210" multiplicity="false">
      <Item>034-911223344</Item>
    </Number>
  </Telephone>
  <Telephone>
    <InternationalCode obligatory="false" code="000.010.000.240" multiplicity="false">
      <Item>034</Item>
    </InternationalCode>
    <Number obligatory="false" code="000.010.000.240" multiplicity="false">
      <Item>611223344</Item>
    </Number>
  </Telephone>
  <Fax>
    <InternationalCode obligatory="false" code="000.010.000.220" multiplicity="false">
      <Item>034</Item>
    </InternationalCode>
    <Number obligatory="false" code="000.010.000.220" multiplicity="false">
      <Item>911223344</Item>
    </Number>
  </Fax>
  <InternetEmailAddress obligatory="true" code="000.010.000.230" multiplicity="true">
    <Item>test@test.com</Item>
  </InternetEmailAddress>
</Contact>
</Agent>
</tns:CVN>

```

Figura 5.5: Campos para el nodo *Agent* (II)

Campos del nodo *CvnItem*

El nodo *CvnItem* se utiliza cuando se hace referencia a un elemento curricular del CVN. Los campos que puede contener este nodo son variables, a diferencia de los dos casos anteriores (*Version* y *Agent*), cuyos campos son fijos. Esto se debe a que cada elemento curricular puede contener distintos tipo de información. Por ejemplo, los campos de un elemento que haga referencia a un congreso diferirán de los de otro elemento que haga referencia a la publicación de un artículo.

Únicamente existe un nodo dentro de *CvnItem* que es fijo, y es conocido

como *CvnItemID*, que identifica el tipo de elemento con un código.

A continuación, se muestran todos los nodos que podría contener el nodo *CvnItem*

- **CvnItemID**: Identifica la categoría a la que pertenece el elemento CVN
- **Subtype**: Especifica de manera más concreta la categoría del elemento
- **Title**: Título del elemento
- **Description**: Descripción del elemento
- **Subject**: Describe la temática que caracteriza al elemento curricular
- **Filter**: Define una lista de elementos que añaden o restringen información
- **Edition**: Permite identificar el número de veces que un elemento es generado
- **Link**: Permite describir aspectos de un elemento curricular heredando las propiedades de otro elemento
- **Entity**: Describe instituciones legales
- **Date**: Descripción de fechas relacionadas con el elemento curricular
- **Author**: Define los participantes del elemento curricular
- **Dedication**: Grado de dedicación de un elemento
- **Roll**: Tipo de responsabilidad del propietario del currículum en relación al elemento
- **Collaborator**: Indica el número de personas involucradas en la ejecución de una determinada actividad
- **Place**: Describe un lugar geográfico para el elemento curricular
- **EconomicDimension**: Ofrece información sobre la dimensión económica del elemento curricular
- **PhysicalDimension**: Ofrece información sobre la extensión o dimensión de un parámetro en el que pueda ser medido el elemento curricular
- **Location**: Permite definir la localización exacta de un elemento curricular
- **ExternalPK**: Registra una secuencia de caracteres usados para identificar un elemento curricular de manera independiente a cualquier otro sistema

- **LocalPK:** Registra una secuencia de caracteres usados para identificar un elemento dentro de algún sistema
- **Scope:** Devuelve información sobre el alcance o dominio geográfico del elemento curricular
- **Diffusion:** Determina canales de difusión alternativos o paralelos a aquellos utilizados tradicionalmente
- **Quality:** Registra información de la codificación de cualquier tipo de información
- **Url:** Guarda una dirección que apunta a un recurso mostrado sobre el elemento curricular
- **Language:** Describe el lenguaje utilizado
- **Validation:** Describe información necesaria para la validación del elemento curricular
- **Contact:** Colecciona información necesaria para contactar a una determinada persona
- **Collection:** Describe un elemento para indicar a que compendio pertenece

A continuación, se muestra un ejemplo en la figura 5.6 del nodo *CvnItem* referente a la categoría de SICA para Aportaciones a Congreso, en donde se muestran algunos de los campos comentados.

Con esto, queda detallada la estructura del formato CVN-XML, describiendo todos los nodos y campos que serán utilizados para obtener la información curricular a la hora de analizar un CVN desde la herramienta que se desarrollará.

```

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<ns2:CVN xmlns:ns2="http://cv.normalized.org/cvn">
  <CvnItem>
    <CvnItemID>
      <CVNPK code="060.010.020.000">
        <Item>060.010.020.000</Item>
      </CVNPK>
      <CodeCvnItem code="060.010.020.000">
        <Item>100</Item>
      </CodeCvnItem>
      <ExperienceType code="060.010.020.000">
        <Item>020</Item>
      </ExperienceType>
      <ResultType code="060.010.020.000">
        <Item>000</Item>
      </ResultType>
    </CvnItemID>
    <Subtype>
      <SubType1 code="060.010.020.010">
        <Item>015</Item>
      </SubType1>
    </Subtype>
    <Title>
      <Name code="060.010.020.030">
        <Item>Video Tester ¿ A Multiple-Metric Framework for Video Quality Assessment over IP Networks</Item>
      </Name>
      <Type>000</Type>
      <Category>000</Category>
    </Title>
    <Link>
      <Link>
      </Link>
    </Link>
    <Author>
      <Author>
      </Author>
    </Author>
    <Author>
      <GivenName code="060.010.020.040">
        <Item>JORGE</Item>
      </GivenName>
      <FirstFamilyName code="060.010.020.040">
        <Item>NAVARRO</Item>
      </FirstFamilyName>
      <SecondFamilyName code="060.010.020.040">
        <Item>ORTIZ</Item>
      </SecondFamilyName>
      <Signature code="060.010.020.040">
        <Item>Navarro-Ortiz, Jorge</Item>
      </Signature>
      <SignatureOrder code="060.010.020.040">
        <Item>2</Item>
      </SignatureOrder>
      <Roll>050</Roll>
    </Author>
  </CvnItem>

```

Figura 5.6: Campos referentes al nodo *CvnItem* para una aportación a congreso

Capítulo 6

Diseño e implementación.

En el presente capítulo se van a abordar los aspectos referentes al diseño y a la implementación de la herramienta que se pretende desarrollar en este proyecto, cuyos términos teóricos se han desarrollado anteriormente.

6.1. Diseño

En esta sección se comentarán los temas referentes al diseño del proyecto, considerando los lenguajes de programación y herramientas utilizadas.

6.1.1. Lenguajes utilizados

Para el desarrollo de la herramienta encargada de la importación de CVN a la plataforma SICA se ha utilizado el lenguaje PHP [19], en su última versión (7.19) que es estable y corrige errores de seguridad de versiones anteriores. PHP es un lenguaje de código abierto adecuado para el desarrollo web y que puede ser incrustado en HTML. Las páginas de PHP contienen HTML con código incrustado que realiza una determinada acción. Para realizar las pruebas necesarias esta herramienta se emulará de manera local, es decir, sin una máquina remota conectada a Internet. Para ello, se hará uso del conocido software XAMPP [20], una distribución de Apache completamente gratuita y que contiene MariaDB, PHP y Perl.

Para la parte del cliente, es decir, lo que se visualizará por pantalla al usuario, se han utilizado los conocidos lenguajes HTML5, JavaScript y CSS3 [21]. El lenguaje HTML aporta la estructura básica de los sitios, mejorada por otras tecnologías como CSS y JavaScript. Es un lenguaje de etiquetas en el que se desarrollan las páginas web. Las hojas de estilo en cascada o CSS se utilizan para controlar la presentación, el formato y el diseño de la aplicación, dando forma al estilo que se desea mostrar. Por último, JavaScript se utiliza para controlar el comportamiento de los diferentes elementos, aportando funcionalidad y dinámica a los componentes de la aplicación.

Finalmente, la herramienta trabajará con el formato XML [22], conocido como Lenguaje de Marcado Extensible. Se trata de un meta-lenguaje que permite definir lenguajes de marcas y que es utilizado para almacenar datos en forma legible. Permite definir la gramática de lenguajes específicos, como HTML, para estructuras documentos. Este lenguaje será utilizado en el formato CVN-XML para almacenar la información del CV, por lo que será necesario trabajar en el análisis del mismo.

6.1.2. Man in the middle proxy (mitmproxy)

Para el análisis del tráfico al generar elementos en SICA se ha utilizado una herramienta conocida como mitmproxy [23], un proxy *man in the middle* interactivo para HTTP y HTTPS con una interfaz por consola. También se ha utilizado otra de sus aplicaciones, conocida como mitmweb, que posea una interfaz basada en web para mitmproxy.

Las características de este software son las siguientes:

- Intercepta peticiones y respuestas HTTP y HTTPS, con la capacidad de modificarlas en el instante
- Guarda conversaciones HTTP completas para su posterior análisis y repetición
- Vuelve a repetir las conversaciones HTTP en el lado del cliente
- Repite las respuestas HTTP de un servidor grabado anteriormente
- Modo *reverse* proxy para reenviar el tráfico a un servidor especificado
- Modo de proxy transparente en OSX y Linux
- Creación de cambios HTTP por *scripting* utilizando Python
- Generación de certificados SSL/TLS utilizados para intercepción en el momento

En la siguiente figura se observa la interfaz utilizada en mitmweb:

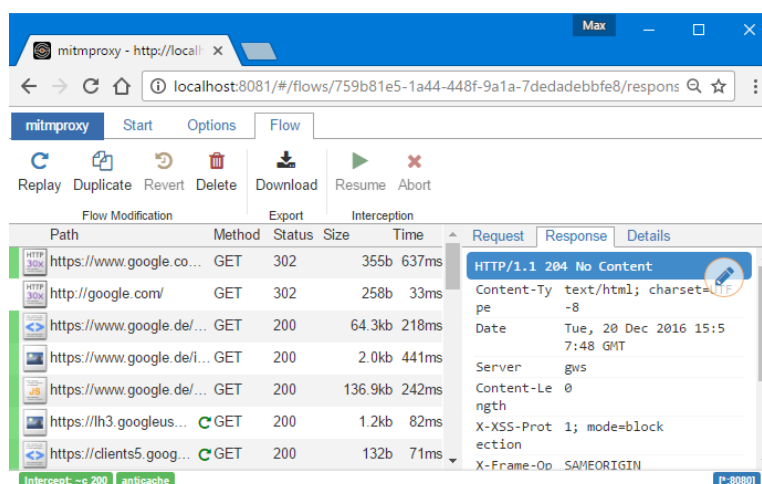


Figura 6.1: Interfaz de mitmweb

Para que mitmproxy funcione correctamente habrá que configurar el cliente como un proxy explícito, de manera que pueda interceptar tráfico. El protocolo del proxy es codificado en el RFC de HTTP, por lo que el comportamiento del cliente y el servidor está bien definido y es fiable. En la interacción más simple posible con mitmproxy, un cliente se conecta directamente al proxy y realiza una petición a un servidor, por lo que en ese momento el proxy reenvía la petición a dicho servidor.

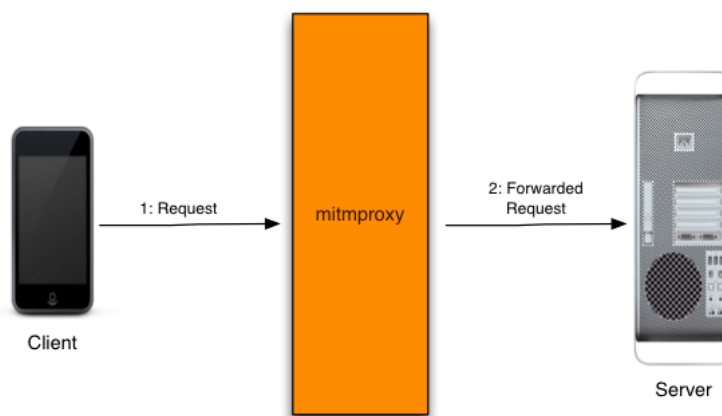


Figura 6.2: Funcionamiento de mitmproxy explícito

Las siglas de MITM (Man-In-The-Middle) hacen referencia al proceso utilizado para interceptar e inferir paquetes de datos teóricamente opacos. La idea básica es pretender ser el servidor para el cliente y pretender ser el cliente para el servidor, mientras que la herramienta se sitúa en el medio de ambos decodificando el tráfico proveniente de ambos lados. Sin embargo, el

sistema de Autoridades Certificadoras (CA) está diseñado para prevenir este tipo de ataques, permitiendo que una tercera parte de confianza firme criptográficamente un certificado del servidor para verificar que es legítimo. Si la firma no coincide o si proviene de una tercera parte que no es de confianza, un cliente seguro cortará la conexión y se negará a proceder con ella. A pesar de las deficiencias del sistema de Autoridades Certificadoras, es muy complicado intentar realizar un *man in the middle* en una conexión TLS para su análisis. Mitmproxy solventa este problema convirtiéndose en una Autoridad Certificadora de confianza por si misma. Para ello, mitmproxy incluye una implementación completa de CA que genera certificados de certificación sobre la marcha. Para que el cliente confíe en estos certificados, se registra mitmproxy como una CA de confianza de forma manual.

Los pasos para realizar este proceso son los siguientes:

1. El cliente realiza una conexión a mitmproxy y emite una petición de conexión HTTP
2. Mitmproxy responde con *200 Connection Established*, como si hubiese configurado la conexión
3. El cliente cree que está conectado con el servidor remoto e inicia la conexión TLS. Utiliza SNI (Server Name Indication) para indicar el nombre del servidor al que se está conectando
4. Mitmproxy conecta al servidor y establece una conexión TLS utilizando el nombre del servidor SNI indicado por el cliente
5. El servidor responde con el certificado correspondiente, que contienen los valores del CN (Common Name) y SAN (Subject Alternate Name) necesarios para generar el certificado.
6. Mitmproxy genera el certificado y continua en el proceso de conectarse a través de TLS, que quedó pausado en el paso 3.
7. El cliente envía la petición sobre la conexión TLS establecida.
8. Mitmproxy pasa la petición al servidor a través de la conexión TLS iniciada en el paso 4.

En la siguiente figura puede observarse el proceso de conexión y generación de certificados:

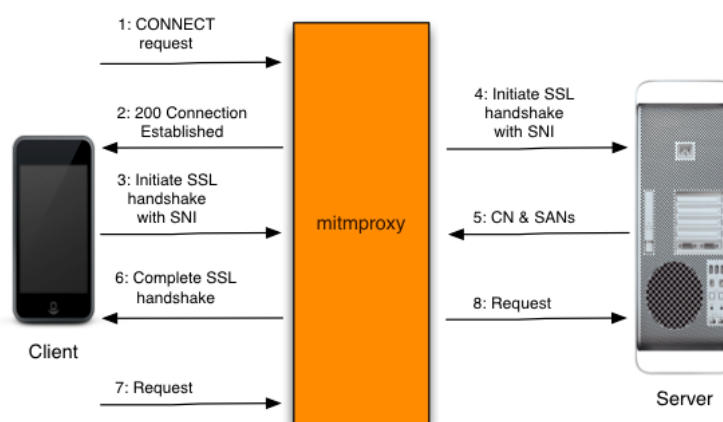


Figura 6.3: Conexión y creación de certificados utilizando HTTPS.

Por lo tanto, con mitmproxy es posible observar el tráfico generado mediante una conexión HTTPS, lo que servirá para analizar el tráfico entrante y saliente al generar elementos en la plataforma SICA, consiguiendo obtener toda la información necesaria para el proceso de importación de CVN a SICA.

Una vez conocidos los fundamentos teóricos de mitmproxy se detalla la arquitectura utilizada para obtener tráfico de SICA. En el caso de este proyecto se ha utilizado mitmweb, una versión web de mitmproxy. Para ello, se ejecuta mitmweb de manera local especificando un puerto o utilizando el establecido por defecto (8080), lo que iniciará un servidor proxy en dicho puerto y una interfaz web en el puerto 8081. Por lo tanto, habrá que configurar el navegador para que utilice como proxy la dirección local 127.0.0.1:8080. Una vez configurado el navegador, se podrá realizar una petición a cualquier página web y en la interfaz web se registrarán todas aquellas direcciones visitadas, tal y como se observa en la figura 6.4. Por lo tanto, todas las peticiones que se realicen a cualquier dirección web mientras se está utilizando mitmproxy serán reenviadas en primer lugar a este proxy, que seguidamente reenviará la información al servicio al que se intentó conectar.

Path	Method	Status	Size	Time
https://sica2.oica.es/static/images/repoblan.svg	GET	200	2.1kb	472ms
https://sica2.oica.es/	GET	200	8.3kb	275ms
https://sica2.oica.es/static/css/home-arc.css?v2.0.1-SNAPSHOT	GET	200	15.7kb	646ms
https://sica2.oica.es/static/js/jquery-1.11.1-SNAPSHOT	GET	200	75.8kb	851ms
https://sica2.oica.es/static/images/logo_sica_registered.png?v2.0.1-SNAPSHOT	GET	200	10.1kb	1s
https://sica2.oica.es/static/fonts/Mul-Regular-webfont.woff	GET	200	24.6kb	1s
https://sica2.oica.es/static/images/contacto.svg	GET	200	2.1kb	433ms
https://sica2.oica.es/static/css/font-guadalupe.css?v2.0.1-SNAPSHOT	GET	200	30.9kb	427ms
https://sica2.oica.es/static/images/logo_jurita.png	GET	200	15.7kb	480ms
https://sica2.oica.es/static/images/bg.png?v2.0.1-SNAPSHOT	GET	200	1.1mb	760ms
https://sica2.oica.es/static/images/logo.png	GET	200	4.7kb	2s
https://sica2.oica.es/investigan/Logout?_1532874182293=	GET	200	0	735ms
https://sica2.oica.es/investigan/Logout?_1532874182293=	GET	404	3.8kb	726ms
https://sica2.oica.es/static/images/favicon.ico	GET	200	1.4kb	940ms

Figura 6.4: Interfaz de mitmproxy web.

Una vez realizada la petición, la respuesta del servidor también pasará por el proxy, por lo que quedará registrada la información, y posteriormente será enviada al equipo cliente.

6.2. Implementación

En esta sección se detallará la implementación que se ha llevado a cabo para la realización de la herramienta. Esta implementación se compone de distintos pasos, entre los que se encuentran el análisis del formato CVN, el análisis del tráfico de la plataforma SICA a la hora de añadir todos los posibles *items* y el desarrollo mediante programación de la herramienta que permitirá importar archivos en formato CVN a la plataforma SICA.

6.2.1. Análisis del formato CVN

En esta subsección se explicará cómo se ha realizado el análisis del formato CVN-XML. Como ya se ha mencionado en la sección 5, un CVN es un archivo PDF con un XML incrustado, en el cual existen diversos campos que especifican el contenido del Currículum Vitae.

Para realizar esta tarea se obtuvo inicialmente un fichero CVN-PDF con diversos elementos de un CV, como pudieran ser la situación profesional, participación en congresos o artículos publicados. En esta primera toma de contacto se investigó como se podía obtener el XML incrustado en el PDF, información que parece ser solo es proporcionada a las entidades que están certificadas por FECYT. En la figura 6.5 se puede observar un ejemplo de cómo se visualiza un archivo PDF abriéndolo con cualquier visualizador de dicho formato:

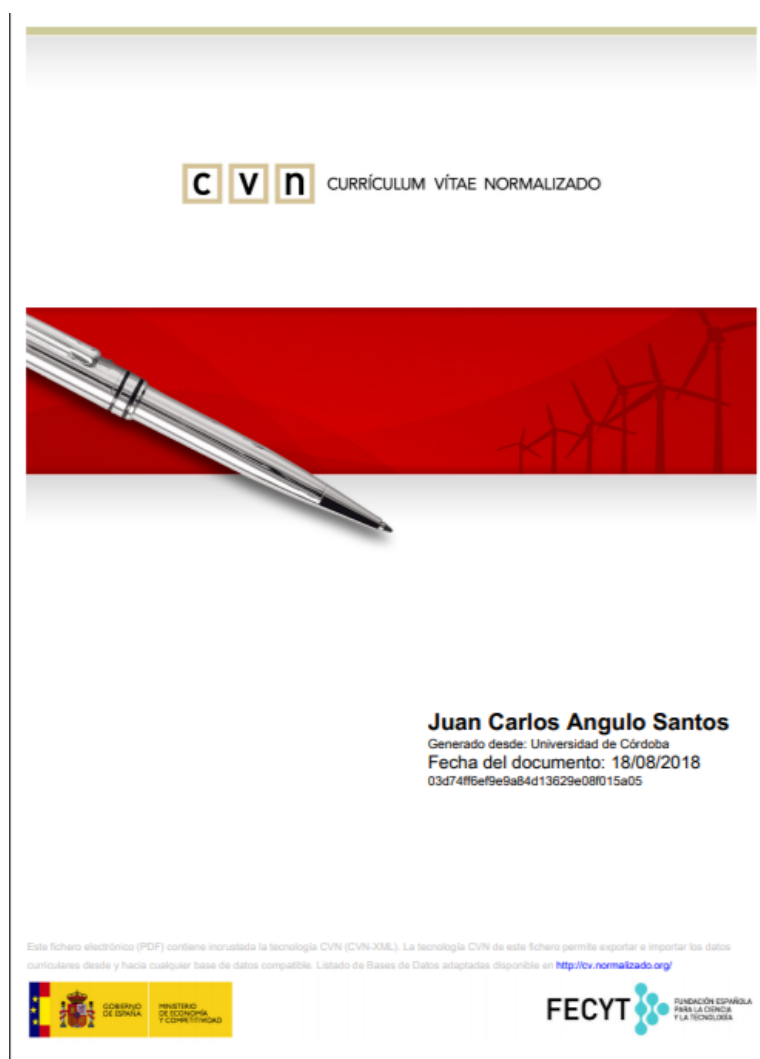


Figura 6.5: Ejemplo de CVN-PDF abierto con visualizador de PDF

Sin embargo, es posible abrir un fichero PDF con editor de textos. Al hacer esto con un fichero CVN-PDF se pueden observar un conjunto de caracteres alfanuméricos, tal y como se observa en la figura



Figura 6.7: Código Base64 incluido en el CVN-PDF

Una vez obtenido dicho código en Base64 se pasó a decodificarlo. El resultado de este proceso devuelve un archivo comprimido de tipo ZIP con nombre *cv.zip* y que en su interior contiene un fichero CVN-XML con nombre *cv.xml*. En este archivo XML se encuentran los datos del CV, tal y como se observaba en la figura 5.1.

Por lo tanto, una vez que se ha obtenido el fichero CVN-XML es posible analizar todos los componentes y estructura del archivo. En primer lugar,

era necesario identificar los nodos del CVN-XML más importantes y conocer a que hacen referencia. Como se comentó en el capítulo 5, existen tres nodos principales (Version, Agent y CvnItem) de los que cuelga información de diferentes tipos. De esos tres nodos, el CvnItem es el que contiene la información curricular, y existe un nodo de ese tipo por cada elemento que contenga el CV. Por lo tanto, ese es el nodo que interesa para la realización de este proyecto, ya que ahí se encuentran los datos del CV que hacen referencia, por ejemplo, al título del *item* o a la categoría a la que pertenece.

Analizando el nodo CvnItem se comprueba que existe un campo cuyo nombre es CVNPK, incluido en el nodo CvnItemID, que contiene un identificador del tipo de elemento incluido en el CV. Puede observarse esto en la figura 6.8. Este identificador hace referencia a un *item* que pertenece a una determinada categoría en el editor de FECYT. Por ejemplo, el código identificativo 060.010.020.000 hace referencia a la sección 'Actividad científica y tecnológica' y subsección 'Trabajos presentados en congresos nacionales o internacionales'.

```
<CvnItem>
  <CvnItemID>
    <CVNPK code="060.010.010.000">
      <Item>060.010.010.000</Item>
    </CVNPK>
    <CodeCvnItem code="060.010.010.000">
      <Item>080</Item>
    </CodeCvnItem>
    <ExperienceType code="060.010.010.000">
      <Item>020</Item>
    </ExperienceType>
    <ResultType code="060.010.010.000">
      <Item>000</Item>
    </ResultType>
  </CvnItemID>
</CvnItem>
```

Figura 6.8: Identificador del tipo de elemento en el CV (CNPK).

Sabiendo esto, se analizaron los distintos campos que puede incluir cada tipo de elemento según su identificador, como por ejemplo lugares, fechas o entidades, y se concluyó que los nodos que puede incluir un nodo del tipo *CvnItem* varían según el tipo de elemento. Por lo tanto, para conocer todos los posibles nodos que puede incluir cada tipo de elemento según su identificador ha sido necesario generar todos los posibles elementos existentes en SICA, completando todos los campos posibles y descargando el CV en formato CVN-PDF posteriormente. Tras esto, se analiza el CVN-XML incrustado y se observan todos los posibles nodos que se incluyen para ese tipo de elemento. Por ejemplo, en la figura 6.9 se observa que para el tipo de elemento 060.010.010.000 existe un nodo que hace referencia al título del elemento y otro a la entidad del mismo.

```

<Title>
  <Name code="060.010.010.030">
    <Item>Nuevas tecnologías en Internet</Item>
  </Name>
  <Type>000</Type>
</Title>
<Entity>
  <EntityName code="060.010.010.100">
    <Item>Universidad de Granada</Item>
  </EntityName>
</Entity>

```

Figura 6.9: Ejemplo de nodos de información en CVN-XML

Repitiendo esto para todos los posibles elementos que pueden ser generados en SICA se ha realizado el análisis del formato CVN y de todos los posibles campos que puede contener para cada tipo de elemento. Cabe destacar que existen ocasiones en el que un elemento de SICA contiene información en su página web que no se ve reflejada en el CVN-XML, ya que no existen nodos para contener esa información o SICA ha decidido no incluirlos al exportar el CV al formato CVN.

6.2.2. Análisis de tráfico en SICA

Para conseguir añadir elementos en SICA a partir de la herramienta que se programará para este proyecto será necesario analizar el tráfico que se genera al introducir un elemento en la página web de la plataforma. Previamente al análisis del tráfico es necesario analizar cómo está estructurada la plataforma SICA y qué elementos pueden ser añadidos en ella. Para ello, es necesario acceder a la página web <https://sica2.cica.es/>.



Figura 6.10: Página de inicio de la plataforma SICA

En la figura 6.10 se observa la página inicial de SICA tras haber iniciado sesión. En esta página se visualizan cinco tipos de contenido que es posible añadir a la plataforma. Estos tipos de contenido son los siguientes:

- Publicación en revista
- Capítulo de libro
- Libro
- Aportaciones a congreso
- Otros

Cada uno de estos tipos de contenido tienen su forma particular de ser añadidos en SICA, por lo que no se sigue un patrón estándar y puede ser necesario acceder a dos o tres páginas más para completar la generación de un nuevo elemento curricular. Los cuatro primeros tipos que se han nombrado permiten añadir un elemento curricular de ese mismo tipo, siguiendo cada uno una estructura determinada, siempre con los mismos campos. Sin embargo, la categoría Otros varía respecto a las anteriores, ya que permite seleccionar entre un numeroso conjunto de tipos de elementos y subtipos que añadir, tal y como se observa en la figura 6.11, y cada uno de ellos sigue una estructura diferente, cambiando continuamente los campos a mostrar.

Inicio / Investig-AN / Mi Currículum

Contacto Español English

AÑADIR ÍTEM - Otros

Tipo de ítem *

Docencia impartida

Seleccione un tipo de ítem...

Actividades genéricas docentes

Actividades genéricas sanitarias

Becas y Reconocimientos

Cargos y Actividades Profesional Anteriores

Catálogo de Obra Artística

Colaboraciones con Centros I+D

Comités científicos de revistas

Comités, Foros, Redes y Sociedades

Conferencias Impartidas

Congresos, Exposiciones, etc.

Convenios y Contratos

Edición de Actas de Congresos, etc.

Edición de Revistas

Ensayos Clínicos

Entrevistas en Medios de Comunicación

Estancias

Experiencia en Evaluación

Formación Universitaria

Subtipo

Docencia impartida

Seleccione un tipo de ítem...

Actividades genéricas docentes

Actividades genéricas sanitarias

Becas y Reconocimientos

Cargos y Actividades Profesional Anteriores

Catálogo de Obra Artística

Colaboraciones con Centros I+D

Comités científicos de revistas

Comités, Foros, Redes y Sociedades

Conferencias Impartidas

Congresos, Exposiciones, etc.

Convenios y Contratos

Edición de Actas de Congresos, etc.

Edición de Revistas

Ensayos Clínicos

Entrevistas en Medios de Comunicación

Estancias

Experiencia en Evaluación

Formación Universitaria

Buscar *

Continuar

Cancelar

Andalucía | R.G.P.D. | Manual | FAQs

Figura 6.11: Página de añadir elemento de categoría Otros en SICA

La categoría Otros es sin duda la más complicada de analizar, puesto que contiene numerosos elementos a generar y cada uno hace uso de campos diferentes, sin seguir una estructura uniforme. Esta categoría únicamente hace uso de dos páginas para su generación, una es la vista en la figura 6.11 y la otra es la pantalla en donde se muestran todos los campos a rellenar, mostrada en la figura 6.12.

DOCENCIA INTERNACIONAL

Área de tipología **Área de título y descripción** Rol en el ítem curricular

Tipo de asignatura

Asignatura

Tipo de docencia

Centro docente

Número de horas/créditos impartidos por curso

Modalidad

Ciclo

Curso

Fecha inicio

Fecha fin

Lugar

Figura 6.12: Página final para añadir elementos de categoría Otros en SICA

A continuación, se muestra un diagrama en la figura 6.13 en el que se muestra el proceso a seguir para añadir un elemento a SICA según la estructura comentada.

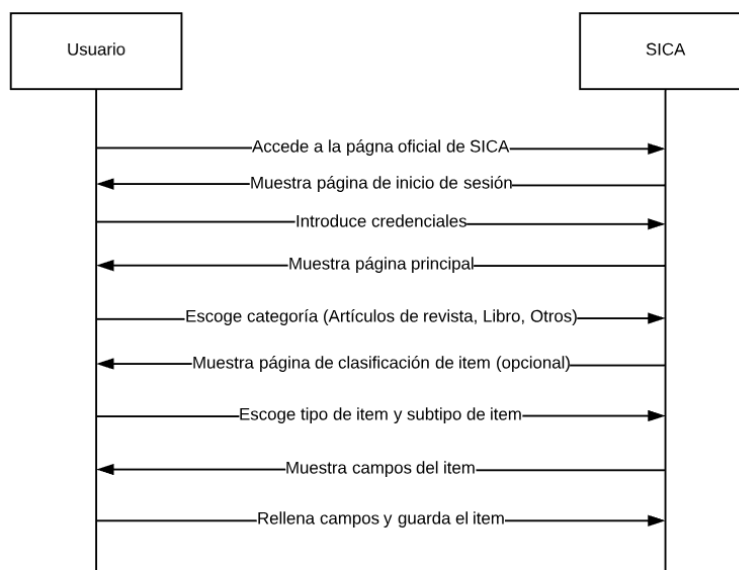


Figura 6.13: Diagrama de pasos a seguir para añadir elemento a SICA

Una vez conocida la estructura de la página web de SICA y como se añaden elementos, se procedió al análisis del tráfico, que se ha realizado con la herramienta mitmproxy, explicada en 6.1.2, que permitirá analizar todas las peticiones enviadas desde un navegador y las respuestas recibidas desde el servidor. En las peticiones que se envíen se podrán observar el nombre y el valor de todas las variables que son enviadas a través de un formulario y a qué página web, y en las respuestas se podrá observar el código HTML que se recibe tras la petición. En la figura 6.14 se observa un ejemplo de cómo se ve una petición utilizando mitmproxy en su versión web.

Request	Response	Details
POST https://sica2.cica.es/investigacion/pages/workforms/newsWorkforms/fatherNewItem.jsf HTTP/1.1		
Host	sica2.cica.es	
Connection	keep-alive	
Content-Length	988	
Origin	https://sica2.cica.es	
User-Agent	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit	
Content-Type	application/x-www-form-urlencoded; charset=UTF-8	
Accept	*/	
Referer	https://sica2.cica.es/investigacion/pages/workforms/news	
Accept-Encoding	gzip, deflate, br	
Accept-Language	es-ES,es;q=0.9,en;q=0.8	
Cookie	JSESSIONID=yRtyFl7w1Z5ZRGpgfgw1GwCwQ0NexelAvoCoKHpE.i xtmmXNaRmtf-.josso02; JOSO_SESSIONID_josso=B54FA26A	
AJAXREQUEST: _viewRoot		
formFatherNewItem: j_id92: level3		
formFatherNewItem: subs2: 1		
formFatherNewItem: aneparea: org.jboss.seam.ui.NoSelectionConverter.noSelectionValue		
formFatherNewItem: titleLB21: Datos CV		
formFatherNewItem: calendarLB22InputDate:		
formFatherNewItem: calendarLB22InputCurrentDate: 08/2018		
formFatherNewItem: calendarLB23InputDate:		
formFatherNewItem: calendarLB23InputCurrentDate: 08/2018		
formFatherNewItem: descriptionLB25:		
formFatherNewItem: textPlacelB26:		
formFatherNewItem: suggPlacelB26_selection:		
formFatherNewItem: entityLB32:		
formFatherNewItem: suggestionEntityLB32_selection:		
formFatherNewItem: j_id327: org.jboss.seam.ui.NoSelectionConverter.noSelectionValue		
formFatherNewItem: formFatherNewItem		
autoScroll:		
javax.faces.ViewState: j_id6		
formFatherNewItem: authTableSharingZero:0:editItemLink:formFatherNewItem:authTableSharingZero:0:editItemLink		
ajaxSingle: formFatherNewItem:authTableSharingZero:0:editItemLink		
AJAX:EVENTS_COUNT: 1		

Figura 6.14: Ejemplo de petición HTTPS en SICA visualizado en mitmweb

Por otro lado, en la figura 6.15 se observa cómo se visualiza en mitmweb una respuesta típica devuelta por SICA, en donde se aprecia el código HTML que se cargará en el navegador.

Request	Response	Details
HTTP/1.1 200 OK		
Server	nginx	
Date	Sat, 18 Aug 2018 17:48:47 GMT	
Content-Type	text/xml; charset=UTF-8	
Content-Length	4950	
Connection	keep-alive	
Expires	0	
Cache-Control	no-cache, must-revalidate, max_age=0, no-store	
X-Powered-By	Undertow/1	
X-Powered-By	JSF/1.2	
X-UA-Compatible	IE=EmulateIE8	
Pragma	no-cache	
Ajax-Response	true	
<pre><?xml version="1.0"?> <html lang="es" xmlns="http://www.w3.org/1999/xhtml"> <head> <title></title> </head> <body> <div id="formFatherNewItem:gridAuthorForm" class="ancho100p"> <div class="ancho100p dark-box-margin"> <div class="ancho100p" style="margin-bottom:15px;"> Datos del autor <img alt="" class="rich-spacer" height="1" id="formFatherNewItem:j_id180" src="/investigan/a4j/g/3_3_3.Finalimages/s Identifique las personas con alguna responsabilidad en el item curricu <script type="text/javascript" id="scriptformFatherNewItem:j_id181"> // <![CDATA[new Tooltip("formFatherNewItem:j_id181","formFatherNewItem:helpL831",{showEvent:'mouseover'}); //]]> </script> <div id="formFatherNewItem:j_id182" style="font-size: 0px;"> <div class="rich-separator ancho95p linea" style="background-image: url(/investigan/a4j/g/3_3_3.Finalong.richfaces </div></pre>		

Figura 6.15: Ejemplo de respuesta desde SICA visualizada en mitmweb

Por lo tanto, conociendo cómo se analiza el tráfico se procede a explicar los pasos que se han llevado a cabo. El proceso de abrir y configurar mitmproxy se explicará en el capítulo 7. Una vez se tiene mitmproxy en su versión web funcionando es necesario iniciar sesión en SICA y ver qué campos se envían en dicho formulario. Esto puede observarse en la figura 6.16, en donde se visualiza el nombre de los campos que se envían y sus respectivos valores. Esto requiere que en la herramienta que se desarrollará se deberá crear una pantalla para iniciar sesión con las credenciales de SICA, ya que es necesario enviarlas en sus correspondientes campos y la dirección especificada en la figura.

POST https://sica2.cica.es/login.do HTTP/1.1	
Host	sica2.cica.es
Connection	keep-alive
Content-Length	136
Cache-Control	max-age=0
Origin	https://sica2.cica.es
Upgrade-Insecure-Requests	1
Content-Type	application/x-www-form-urlencoded
User-Agent	Mozilla/5.0 (Windows NT
Accept	text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Referer	https://sica2.cica.es/
Accept-Encoding	gzip, deflate, br
Accept-Language	es-ES,es;q=0.9,en;q=0.8
Cookie	JOS50_SESSIONID_josso=AlmmXNaRmtf-.josso02

josso_cmd:	login
josso_back_to:	
default_user:	Nombre de usuario
josso_username:	usuario_enviado
default_pwd:	Contraseña
josso_password:	contraseña_enviada

Figura 6.16: Campos enviados en el login de SICA

Este paso se deberá realizar durante todo el proceso de añadir un elemento a SICA, analizando todas las peticiones realizadas. En el diagrama de la figura 6.17 se muestra el proceso de análisis de tráfico con mitmproxy.

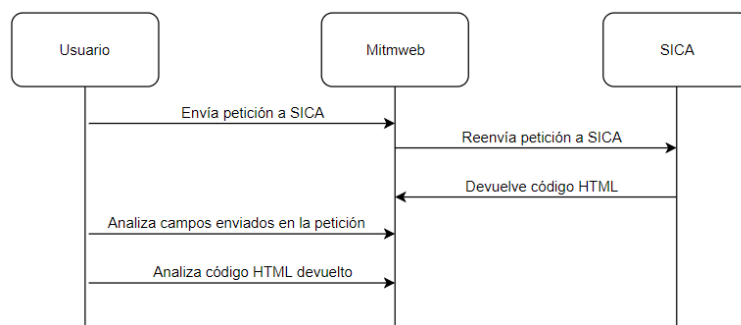


Figura 6.17: Diagrama que muestra el análisis del tráfico con mitmweb en SICA

Los pasos mostrados en el diagrama 6.17 se realizarán las veces necesarias hasta completar la subida de un *item* a SICA.

Una vez iniciada sesión en la plataforma será necesario escoger uno de los cinco tipos de categorías nombrados anteriormente. Esto no es difícil, ya que no hay que enviar ningún tipo de parámetro en la petición y únicamente será necesario conocer la dirección URL para insertar cada uno de los cinco tipos de categorías.

Tras seleccionar el tipo de categoría a importar cambiará el modo de proceder según la escogida, ya que cada una tiene sus propias características de generación como se comentó anteriormente. Será necesario, por lo

tanto, analizar el tráfico de generación todos los posibles tipos y subtipos de elementos de cada categoría para llevar a cabo la finalización de esta tarea, que se utilizará para el desarrollo de la herramienta.

El guardado de elementos en SICA está basado en el uso de formularios (*form*) de HTML, que utilizan los llamados *inputs* o entradas para el envío de los datos del formulario. Este proceso de subida de datos se realiza al pulsar sobre un botón de tipo *submit* que enviará los datos a una URL definida en el campo *action* del formulario. Además, estos datos se envían como un conjunto de datos del tipo clave-valor, siendo la clave el nombre dado al elemento *input* y el valor lo que haya escrito o seleccionado en ese momento en dicho campo del formulario.

Esta información aparece utilizando mitmproxy, como se veía en la figura 6.14, en donde se ven los campos que se han enviado y que deberán utilizarse en la petición para añadir el *item* a SICA. Sin embargo, existen una serie de campos que no tienen un nombre o valor estático, y que pueden cambiar en el caso de que SICA modifique su página, lo que provoca cierta incertidumbre a la hora de saber cuándo cambiaran.

Es el caso de los siguientes campos y valores, que se envían en la petición con el siguiente nombre:

- **javax.face.ViewState:** Este campo es un valor de la forma *j_id_numero*, por ejemplo, *j_id_6* que varía al cambiar la página en la que se encuentra el usuario. Por lo tanto, será necesario encontrar ese valor cada vez que se cambia de página.
- **form:j_id_numero:** Estos campos o valores son entradas del formulario que representan, por lo general, elementos de tipo *select*, botones, o campos que requieren hacer uso de AJAX previamente al envío del formulario. Están formados por la palabra *form* o *formFatherNewItem*, seguidos de dos puntos, más los caracteres *j_id* y un número entero, por ejemplo. Estos campos son estáticos por lo general, pero el número entero puede cambiar si los desarrolladores de SICA introducen nuevos cambios en la página web. En la figura 6.18 puede observarse un código HTML con uno de estos campos.

```
▼<div id="formFatherNewItem:scopeLB41" class="ancho100p">
  ▼<select name="formFatherNewItem:j_id339" class="ancho45p" size="1"> %0
    <option value="org.jboss.seam.ui.NoSelectionConverter.noSelectionValue"
      selected="selected">Seleccione un alcance...</option>
    <option value="27">Autonómica</option>
    <option value="28">Internacional no UE</option>
    <option value="29">Nacional</option>
    <option value="30">Unión Europea</option>
  </select>
```

Figura 6.18: Campos del tipo form:j_id

Por lo tanto, para encontrar este tipo de valores dinámicos en los que por lo general cambia un número, será necesario utilizar expresiones regulares para encontrar el valor en el código HTML de la página que se está visitando en un determinado momento. Las expresiones regulares [25] son un método por medio del cual se pueden realizar búsquedas dentro de cadenas de caracteres, sin importar la amplitud de búsqueda requerida de un patrón definido de caracteres, proporcionando una solución práctica para encontrar lo deseado dentro de una cadena de texto.

Cabe destacar, que para encontrar aquellos campos que requieren de expresiones regulares es necesario proporcionar el mínimo código HTML como sea posible para encontrar el valor deseado. Esto se realiza proporcionando el código HTML que se encuentra delante y detrás del valor buscado. Sin embargo, si la página de SICA cambiara su código HTML estas búsquedas a través de expresiones regulares no servirían, por lo que es importante que el código que se proporcione a la expresión sea poco propenso a cambios, aunque a veces no es posible cumplir esta condición. Para encontrar estos campos dinámicos se proporcionará el código HTML que se encuentra antes y después del campo buscado. Por ejemplo, en la figura 6.18 el campo dinámico buscado es *j_id339*. Este número puede variar en algún momento, por lo que se buscará proporcionando el código HTML que se encuentra antes, es decir, *name="formFatherNewItem:* y el código que se encuentra después, que son simplemente unas comillas dobles (").

En el siguiente diagrama 6.19 se muestra la manera en la que se ha implementado la búsqueda de estos campos dinámicos a través de expresiones regulares:

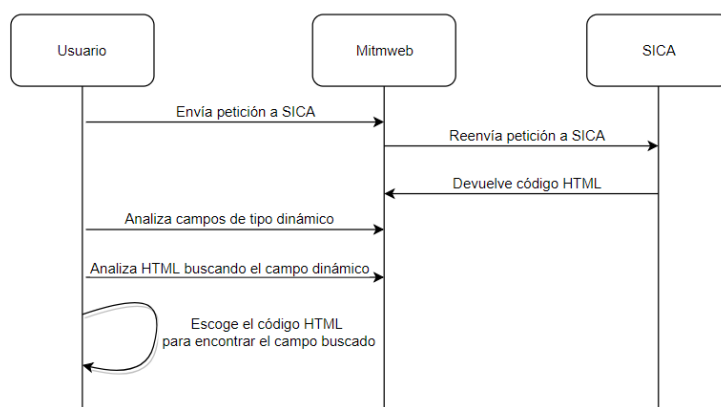


Figura 6.19: Diagrama que muestra como encontrar campos de tipo dinámico

Otro factor importante en el análisis del tráfico son las peticiones AJAX [26]. AJAX, cuyas siglas en inglés significan *Asynchronous JavaScript And XML* o JavaScript asíncrono y XML, es una técnica de desarrollo web para la creación de aplicaciones interactivas. Estas aplicaciones se ejecutan en el

lado del cliente, es decir, en el navegador de los usuarios mientras se mantiene una comunicación asíncrona con el servidor en segundo plano. Muchos de los campos de los formularios en SICA hacen uso de esta tecnología para proporcionar sugerencias de palabras o guardar algunos de los valores escritos o seleccionados por el usuario antes de realizar la subida definitiva del formulario que produce el guardado del elemento. Por lo tanto, es importante tener en cuenta que, además de los campos que se envían al pulsar el botón de tipo *submit* del formulario, también se debe realizar un análisis del tráfico de aquellas peticiones AJAX que poseen algunos campos. Por lo general, los campos que usan AJAX son aquellos en los que se realizan sugerencias al usuario, como, por ejemplo, de lugares o entidades. En la figura 6.20 se encuentra un ejemplo de código HTML en el que se muestra un campo que utiliza una petición AJAX para mostrar sugerencias de palabras cada vez que se escribe en el mismo. Es identificable ya que aparece en el atributo *onchange* la función que realiza la petición (*A4J.AJAX.Submit*) y que se activará cada vez que cambie el valor del campo.

```
<input id="formFatherNewItem:entityLB24"
type="text" name="formFatherNewItem:
entityLB24" value class="ancho95p"
onchange=
"A4J.AJAX.Submit('formFatherNewItem',event
,{ 'similarityGroupId':
'formFatherNewItem:j_id152', 'parameters':
{ 'formFatherNewItem:j_id152':
'formFatherNewItem:j_id152' } } )"
autocomplete="off">
```

Figura 6.20: Campos de formulario con petición AJAX

Con esto, queda explicado el análisis del tráfico, tarea que terminará al analizar todos los elementos que pueden ser añadidos en SICA y que permitirá realizar la importación de elementos.

6.2.3. Desarrollo de la herramienta

En esta sección se explicará cómo se ha desarrollado la herramienta, teniendo en cuenta el contenido de las secciones vistas anteriormente, es decir, el análisis del formato CVN y el análisis del tráfico en SICA.

Como se comentó en el capítulo de diseño, se ha utilizado el lenguaje PHP para el desarrollo de la herramienta de este proyecto, lo que quiere decir que la herramienta estará basada en tecnología web, teniendo una parte cliente que visualizará el usuario, y una parte servidora que se encargará de realizar todo el procesamiento necesario para importar un CVN a SICA. En

esta sección se llevará a cabo la explicación de cómo se han realizado ambas partes.

Lado del cliente

En el lado del cliente se visualizarán las páginas necesarias para que el usuario tome las decisiones que desee realizar con la herramienta. Para su desarrollo se utilizará la tecnología HTML, JavaScript y CSS, tal y como se comentó en la sección de diseño. Adicionalmente, se ha utilizado Bootstrap [27], un conjunto de herramientas de código abierto para diseño de sitios y aplicaciones web. Contiene plantillas de diseño con tipografía, formularios, botones y otros elementos de diseño basado en HTML y CSS, así como extensiones JavaScript. Una de sus principales características es que soporta diseños web adaptables, lo que significa que la página se ajusta dinámicamente, tomando en cuenta las características y tamaño de la pantalla del dispositivo usado. La herramienta desarrollada tiene esa característica, por lo que puede ser utilizada en cualquier tipo de dispositivo sin que el diseño se vea afectado.

En total, existen cuatro páginas que serán visibles por el usuario, es decir, que contengan código HTML. En la tabla 6.1 puede verse resumidamente la dirección URL del dominio para visualizar cada una de las páginas y una breve descripción de las mismas.

URL	Descripción
login	Inicio de sesión en SICA y en la herramienta
logout	Cierre de sesión en la herramienta
importar	Importación de fichero CVN
preseleccion	Resolución de conflictos entre CVN y SICA
seleccion	Selección de elementos a importar

Tabla 6.1: Páginas existentes en la herramienta.

Todas estas páginas serán explicadas a continuación, dando una descripción detallada de cada una y mostrando su funcionamiento.

En primer lugar, será necesaria una pantalla para introducir las credenciales de SICA, ya que no es posible añadir elementos en la plataforma sin iniciar sesión. Esta pantalla de inicio de sesión será bastante simple, con un campo para introducir usuario y otro para introducir la contraseña, además de un botón que permitirá iniciar sesión, tanto en la herramienta como en la plataforma. Cabe destacar, que la herramienta no podrá usarse a menos que el usuario haya iniciado sesión, lo que estará controlado en el servidor. En la figura 6.21 puede observarse como es la página de inicio de sesión visualizada desde un ordenador. Por otro lado, esta será la página inicial de la aplicación y además, podrá visitarse a través de la URL de dominio /login.



INTRODUZCA SUS
CREDENCIALES DE SICA

USUARIO

Test

CONTRASEÑA

.....

Login

Figura 6.21: Página de inicio de sesión en la herramienta

Para mostrar las características adaptables a la pantalla según el dispositivo de la herramienta se puede ver la figura 6.22, en donde se muestra cómo se vería la página de inicio de sesión si se utilizará con un dispositivo móvil Samsung Galaxy S5, que tiene una pantalla de 5.1 pulgadas o equivalentemente de 360x640 píxeles.



Figura 6.22: Página de inicio de sesión adaptable en dispositivo móvil

Mientras el usuario tenga la sesión iniciada en la página de la herramienta, aparecerá en la parte superior de la página un mensaje informativo con el nombre del usuario que ha iniciado sesión y un botón para cerrar la misma, tal y como se ve en la figura 6.23.

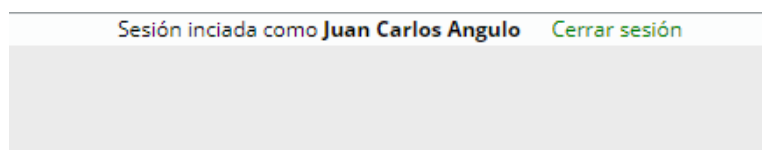


Figura 6.23: Información sobre el usuario que ha iniciado sesión

Al pulsar sobre el botón de cerrar sesión se enviará al usuario a la URL /logout, en donde se realizará el proceso necesario para cerrar la sesión del usuario.

Por otro lado, en caso de que el inicio de sesión no sea exitoso, se mostrará un mensaje de error, tal y como se ve en la figura 6.24.

A login form titled "INTRODUZCA SUS CREDENCIALES DE SICA". It contains two input fields: "USUARIO" and "CONTRASEÑA". The "CONTRASEÑA" field has a toggle icon on the right. Below the fields is a dark "Login" button. A red error message "Autenticación errónea" is displayed below the button. The form is centered on a light gray background.

INTRODUZCA SUS
CREDENCIALES DE SICA

USUARIO

CONTRASEÑA

Login

Autenticación errónea

Figura 6.24: Mensaje de error en la página de inicio de sesión

Una vez que se ha iniciado sesión en la herramienta se redirigirá al usuario a una página para introducir su CVN en la herramienta, cuya URL de dominio es /importar. Esta página es muy sencilla, ya que únicamente hace uso de un botón para seleccionar un fichero CVN-PDF o CVN-XML y otro para comenzar la importación. En la figura 6.25 puede observarse el diseño de esta página.



Figura 6.25: Página para introducir fichero CVN-PDF o CVN-XML

Al pulsar sobre 'Seleccionar archivo' se abrirá una ventana para escoger el fichero, que puede estar alojado en cualquier lugar del equipo del usuario. Esto puede observarse en la figura 6.26. Una vez escogido el fichero aparecerá el nombre del mismo debajo del botón nombrado anteriormente, mostrando la dirección del equipo en la que se encuentra. Esto puede variar dependiendo del navegador, ya que por ejemplo Google Chrome sustituye la dirección del equipo por la palabra *fakepaths* para la prevención de ataques al obtener esa información.

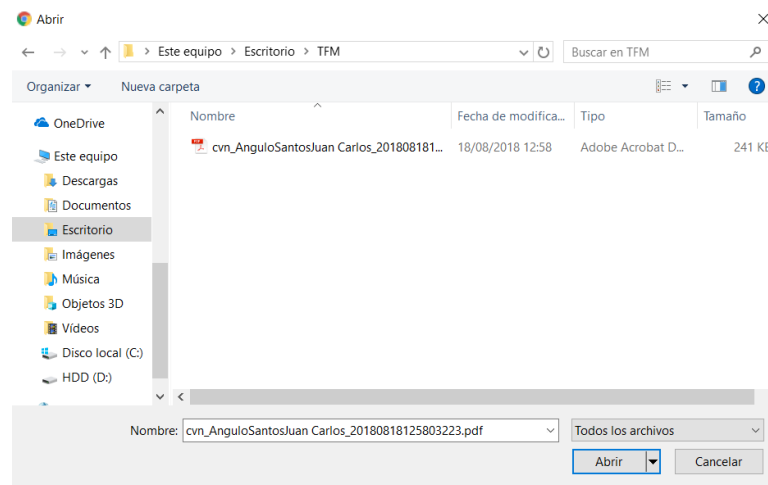


Figura 6.26: Ventana para la selección de fichero

Al pulsar sobre el botón 'Importar' se realizará un procesamiento del fichero que se explicará en la sección del lado del servidor. Este procesamiento incluye mostrar errores en caso de que no se haya subido ningún fichero, no sea un fichero PDF, o no se trate de un fichero del tipo CVN-PDF. En la figura 6.27 se observa un ejemplo de error mostrado en esta página para importar.



Figura 6.27: Error de CVN-PDF no válido en la página de importación

El listado de errores disponibles en esta página son los siguientes:

- No se ha recibido ningún archivo
- No es un archivo PDF o XML
- No es un archivo CVN-PDF válido
- No es un archivo CVN-XML válido

Si se ha subido un fichero CVN-PDF que no contiene errores, es decir, que el archivo *cv.zip* comprimido que se encuentra en su interior es extraíble y además dentro contiene un CVN-XML, o si se ha subido directamente un CVN-XML válido, se pasará a la siguiente página. Esta nueva página puede variar en esta ocasión dependiendo del contenido del CVN-XML, pasando a la página de pre-selección o selección.

Esto es debido a que en el editor de FECYT pueden ser insertados tantos elementos como posibles códigos de identificación puede haber en el formato CVN-XML. Sin embargo, para la plataforma SICA esto no ocurre de esa manera, ya que un mismo código de identificación de *item* del CVN puede corresponder a varias categorías, tipos de elemento o subtipos de elemento de SICA al mismo tiempo. Esto supone que sea necesaria una página para indicar a que categoría, tipo o subtipo de SICA pertenece un elemento encontrado en el CVN.

Por lo tanto, suponiendo que se encuentra un elemento en el CVN que pertenezca a dos o más categorías, tipos o subtipos de elemento de SICA y no pueda ser clasificado en alguno de ellos, se pasará a la página de pre-selección. En esta página se mostrarán los elementos del CVN que presentan conflictos y que deben ser clasificados en una categoría, tipo de elemento o subtipo de elemento. En la figura 6.28 se observa el diseño de esta página, en donde se ha utilizado un elemento de SICA de la categoría Otros, en concreto del tipo de elemento de Ensayos clínicos. El código utilizado para los ensayos clínicos coincide con el utilizado por SICA para el tipo de elemento de Proyectos de I+D+i y Ayudas, por lo que se produce un conflicto que el usuario debe resolver. Para ello, se mostrará el título del elemento encontrado en el CVN, que será la única información que puede ser proporcionada en este tipo de conflictos.

LOS SIGUIENTES OBJETOS PUEDEN PERTENECER A DIFERENTES CATEGORÍAS. ESCOGA DONDE IMPORTARLOS:

Ensayo clínico

Categoría (*)

Otros

Tipo Item (*)

Ensayos clínicos

Ensayos clínicos

Proyectos de I+D+i y Ayudas

Ensayo clínico

Continuar

Figura 6.28: Página de pre-selección para la resolución de conflictos en el CVN y SICA

Como se observa en la figura 6.28, los campos están definidos por defecto para que un elemento del CVN siempre esté clasificado en una categoría, pero debe ser el usuario el que escoja la categoría, tipo de elemento o subtipo de elemento. Lamentablemente, SICA no está preparada para la importación de CVN por la incompatibilidad de los códigos con sus posibles categorías. Sin embargo, la solución propuesta solventa este problema, aunque requiera acciones del usuario.

Tras la resolución de conflictos se deberá pulsar sobre el botón Continuar y se pasará a la siguiente página para seleccionar los elementos del CVN que se desean importar a SICA, la página de selección. Esta página muestra todos los elementos encontrados en el CVN, clasificados en categoría, tipo de elemento y subtipo de elemento de SICA, además de todos los posibles campos que puedan ser requeridos por la plataforma, tanto obligatorios como opcionales.

Esta página se ha realizado con un formulario, cuyos campos de tipo *input* tienen el mismo nombre que los que se pueden encontrar en SICA, con el fin de facilitar la subida de estos elementos. Esto ocurre para aquellos campos con nombre estático, ya que como se comentó en la sección 6.2.2, existen campos con un nombre dinámico que será necesario encontrar a través del propio código HTML de SICA.

En esta página se utiliza el análisis del formato CVN explicado en la sección 6.2.1 para intentar encontrar en el CVN los datos de los campos que aparecen en SICA para un determinado elemento. Sin embargo, muchos de los campos que utiliza SICA no aparecen en el CVN, por lo que si el usuario

lo desea puede rellenarlos en esta página. Los datos referentes a campos de SICA que aparezcan en el CVN serán recogidos y escritos en la página de selección de elementos, evitando así que el usuario tenga que escribirlos nuevamente. Además, se han marcado los campos que son obligatorios para una correcta importación, denegando la importación a no ser que se hayan rellenado todos estos campos.

Una vez se acceda a la página de selección de elementos, aparecerá el título de todos los que se han encontrado en el CVN, indicando entre paréntesis la sección y subsección a la que pertenece dicho código en el editor de FECYT. En un principio estos elementos aparecerán sin desplegar, es decir, sin mostrar sus campos, aunque es posible pinchar sobre el título de cada uno para desplegar la información referente a cada elemento. En la figura 6.29 se puede observar el diseño para esta página.



Figura 6.29: Página de selección de elementos a importar a SICA

Si se pincha sobre el título del elemento se desplegará la información del mismo, mostrando todos sus campos y valores encontrados en el CVN, tal y como se muestra en la figura 6.30. Además, se ha implementado un botón para seleccionar todos los elementos existentes en la página, evitando así que un usuario que desee importar todos o la gran mayoría de elementos encontrados tenga que hacerlo uno por uno.

☒ **Profesor ayudante doctor** (Situación profesional - Cargos y actividades desarrolladas con anterioridad)

Tipo de ítem (*)
Cargos y Actividades Profesional Anteriores ▼

Subtipo (*)
Cargo y/o actividades anteriores a la actual ▼

Título del ÍTEM (*)
Profesor ayudante doctor

Fecha inicio (*)
16/11/2010

Fecha fin
dd/mm/aaaa

Lugar
Granada

Áreas ANEP
Seleccione un área... ▼

Organismo (*)
Universidad de granada

Rol del autor

Rol autor (*)
Becario/a (pre o posdoctoral, otros) ▼

Tipo de dedicación autor
Seleccione tipo de dedicación... ▼

Figura 6.30: Campos de un elemento en la página de selección

Cada elemento encontrado es separado por una línea de color negra, además de diferenciarse por el título en color azul y un *checkbox* (casilla de verificación) que indicará si el elemento será importado a SICA o no. Para seleccionar los elementos que se desean importar deberá marcarse la casilla de verificación, que pasará a estar en un color verde y que indicará que el elemento será importado a SICA.

Al final de la página, una vez que han sido mostrados todos sus elementos y sus respectivos campos, deberá pulsarse un botón con el nombre 'Importar a SICA', que comenzará la importación llamando al servidor encargado del procesamiento. En ese momento aparecerá un símbolo de carga (*loader*) que finalizará una vez se haya completado la importación. Esto puede observarse en la figura 6.31.



Figura 6.31: Final de la página de selección de elementos a importar a SICA

Por otro lado, en caso de que existan campos obligatorios que deban ser completados antes de realizar la importación se mostrará el título del elemento en color rojo, así como se colorearán en rojo los bordes de los campos que deben ser rellenados. Esto se deberá a que SICA considera como obligatorios campos que no existen en el CVN o cuya información no ha podido ser obtenida. Puede observarse esta característica en la figura 6.32.

☒ **Miembro del grupo de investigación**
en TSTC de la UGR (Actividad científica y
tecnológica - Resumen de otros méritos)

Tipo de Ítem (*)
Becas y Reconocimientos ▼

Subtipo (*)
Reconocimiento o Acreditación In ▼

Título del ÍTEM (*)
Miembro del grupo de investigación €

Áreas ANEP
Seleccione un área... ▼

Fecha de concesión (*)
dd/mm/aaaa

Cuantía concedida

Tipo de moneda
EUR ▼

Figura 6.32: Elemento con campos obligatorios sin rellenar

Además, tras realizar el proceso de importación, se mostrará el título de aquellos elementos que no pudieron ser añadidos a SICA debido a un fallo por parte de la herramienta o por parte de SICA. Puede verse un ejemplo

de esto en la figura 6.33.

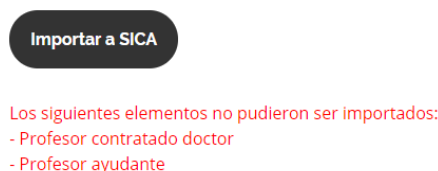


Figura 6.33: Listado de elementos que no pudieron ser importados a SICA

La importación de todos los elementos seleccionados se realiza a través de JavaScript mediante una petición a una página escrita en PHP encargada del procesamiento. Se ha realizado una función en JavaScript para obtener todos los elementos que han sido seleccionados para su importación, creando un JSON en el que se muestra el código de identificación, el valor de todos los campos mostrados, la categoría, tipo de elemento y subtipo de elemento al que pertenece el elemento según su código, aunque estos tres últimos valores del vector se explicarán con más detalle en el lado del servidor. Un JSON [28] es un formato de texto ligero para el intercambio de datos, lo que facilita el análisis por parte del servidor.

En la figura 6.34 se observa un ejemplo del JSON que se enviará al servidor, a través de una petición de tipo POST [29], en la cual se envían los datos a la siguiente página de forma invisible para el usuario.

```
▼ {1: {...}, 2: {...}} ⓘ
▼ 1:
  form:subs2: "41"
  form:title: "Otros"
  form:workformtype: "38"
  formFatherNewItem:aneparea: "2"
  formFatherNewItem:calendarLB22InputCurrentDate: "08/2018"
  formFatherNewItem:calendarLB22InputDate: "12/9/2011"
  formFatherNewItem:calendarLB23InputCurrentDate: "08/2018"
  formFatherNewItem:calendarLB23InputDate: "16/8/2018"
  formFatherNewItem:dedicationTypeId: "38"
  formFatherNewItem:entityLB32: "Universidad de granada"
  formFatherNewItem:roleSelect: "29"
  formFatherNewItem:textPlaceLB24: "Granada"
  identificador: "010.010.000.000"
  index_categoria: "1"
  index_subtipo_item: "1"
  index_tipo_item: "1"
  ▶ __proto__: Object
  ▶ 2: {}
```

Figura 6.34: JSON enviado para el procesamiento de importación a SICA

En la figura 6.35 se muestra un diagrama de flujo resumiendo toda la parte del cliente para entender como se realiza el visualizado de elementos en el lado del cliente, llamando de forma ajena al usuario a páginas PHP que procesarán los datos.

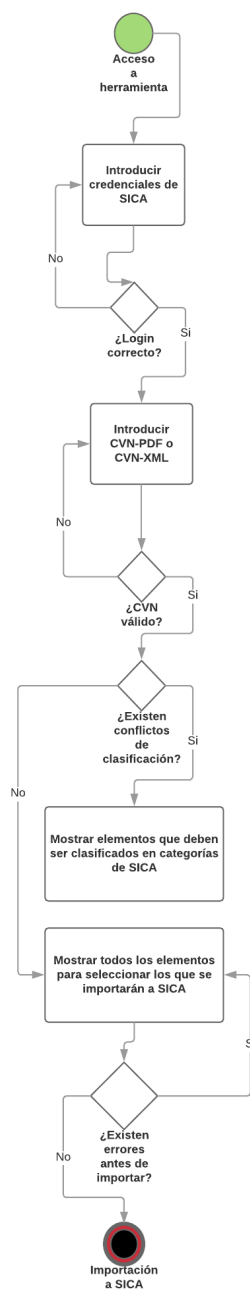


Figura 6.35: Diagrama de flujo del proceso de importación en el lado del cliente

Con esto, queda explicado cómo se ha realizado la parte del cliente, viendo todas las tecnologías utilizadas, el diseño de las diferentes páginas y el modo de proceder en cada una de ellas.

Lado del servidor

En esta sección se explicará cómo se ha desarrollado la parte del lado del servidor, utilizando para ello lenguaje PHP.

En primer lugar, es necesario explicar cómo se han guardado todos los posibles elementos importables a SICA. Esto se ha hecho con utilizando vectores asociativos [30], es decir, un vector con una clave y un valor. Un vector asociativo es un tipo abstracto de dato formado por una colección de claves únicas y una colección de valores, con una asociación uno a uno. En el caso de este proyecto el índice principal del vector será el código de identificación del elemento, y el valor será otro vector asociativo, que a su vez contendrá nuevamente otros índices y valores. Los valores del vector siempre podrán ser otro vector asociativo, una cadena de texto o un número.

Por lo tanto, en el vector en el cual se contiene la información de cada *item* existirán tantos índices como códigos de identificación existan en un CVN, tal y como se observa en la figura 6.36.

```
$tCodeCVNitem1 = [
    //010.010.000.000: Otros -> Situación profesional -> Situación profesional
    "010.010.000.000" => [...],
    //010.020.000.000: Otros -> Cargos y actividades anteriores -> Cargo y/o ac
    "010.020.000.000" => [...],
    //020.010.010.000: Otros -> Formacion univesitaria -> Titulado medio, titu
    "020.010.010.000" => [...],
    //020.010.020.000: Otros -> Formacion universitaria -> Doctor
    "020.010.020.000" => [...],
    //020.060.000.010: Idiomas -> Conocimiento de idiomas
    "020.060.000.000" => [...],
    //020.010.030.000: Otros -> Formación universitaria -> DEA. Máster. Postgre
    "020.010.030.000" => [...],
    //020.020.000.000: Otros -> Otra formación -> Especialidad. Extensión unive
    "020.020.000.000" => [...],
    //020.030.000.000: Otros -> Otra formación -> Curso para la mejora de atenc
    "020.030.000.000" => [...],
    //020.040.000.000: Otros -> Otra formación -> Formación de I+D/Postformacio
    "020.040.000.000" => [...],
    //020.050.000.000: Otros -> Otra formación -> Curso de especialización. Per
    "020.050.000.000" => [...],
    //030.040.000.000: Tesis, Tesinas, DEA, etc. -> Proyecto de fin de carrera.
    "030.040.000.000" => [...],
    //030.010.000.000: Docencia Impartida -> Docencia Internacional. Docencia i
    "030.010.000.000" => [...],
    //No tiene categoria en sica
    "030.020.000.000" => [...],
    //No tiene categoria en sica
    "030.030.000.000" => [...],
];
```

Figura 6.36: Vector con todos los índices con códigos de identificación

Como se observa, se ha añadido a modo de comentario a que categoría

de SICA pertenece cada código de identificación que se muestra como índice del vector. El valor de dicho índice será otro vector asociativo con nuevos índices y valores.

Para aclarar cómo se ha realizado este vector en el cual se guarda la información de todos los elementos existentes en SICA se detallarán varios diagramas en el que se muestra la estructura general, realizando una explicación sobre cada uno de ellos.

Lo visto hasta ahora se puede observar en el diagrama de la figura 6.37, en el que se muestra como el vector principal incluye un número N de códigos identificativos que a su vez determinan un elemento que puede ser importado a SICA.

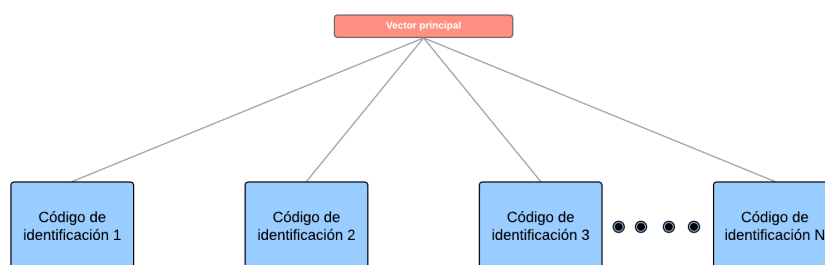


Figura 6.37: Diagrama en el que se muestra el vector principal

Profundizando en un elemento en concreto, se muestra un vector con los siguientes índices: *seccion*, *subseccion*, *titulo* y *categorias*. A continuación, se muestra cada uno y su significado:

- **seccion:** Indica el nombre de la sección del editor de FECYT a la que pertenece dicho código
- **subseccion:** Indica el nombre de la subsección del editor de FECYT a la que pertenece dicho código
- **titulo:** Muestra cómo obtener el título del elemento al buscar en los nodos de un CVN-XML. Se trata de un vector que indica, en orden descendente, todos los nodos por los que habría que pasar para llegar al valor buscado.
- **categorias:** Se trata de un vector que incluye la categoría a la que pertenece el elemento asociado al código principal. Dentro de este vector se encuentra un campo para la clasificación del elemento en un tipo o subtipo, un campo para saber qué elementos HTML deben ser mostrados al usuario y otro con todas las peticiones de tráfico que deben ser enviadas a SICA para la importación del elemento.

En la figura 6.38 se muestra como se estructura el vector del código de identificación en forma de diagrama.

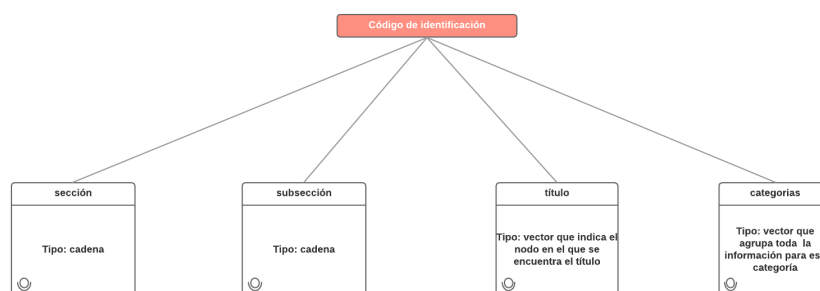


Figura 6.38: Diagrama en el que se muestra el vector del código de identificación

Los valores de *seccion* y *subseccion* que indican cadenas de texto obtenidas de FECYT, se han implementado como variables estáticas en el código, considerando todas las posibles secciones y subsecciones que existen actualmente en el editor FECYT. En la figura 6.39 se puede observar un extracto de las posibles secciones y subsecciones, en forma de variable, que pueden ser encontradas en el editor, y que servirán para identificar elementos a la hora de mostrarlos al usuario.

```

K?php

$seccion2 = "Situación profesional";
$seccion3 = "Formación académica recibida";
$seccion4 = "Actividad docente";
$seccion5 = "Experiencia científica y tecnológica";
$seccion6 = "Actividad científica y tecnológica";
$seccion7 = "Actividad en el campo de la sanidad";
$seccion8 = "Resumen de texto libre";
$subseccion2_1 = "Situación profesional actual";
$subseccion2_2 = "Cargos y actividades desarrolladas con anterioridad";
$subseccion3_1 = "Estudios de 1º y 2º ciclo, y antiguos ciclos (Licenciados, Diplomados, Ingenieros Superiores, Ingenieros Técnicos, Arquitectos)";
$subseccion3_2 = "Doctorados";
$subseccion3_3 = "Conocimiento de idiomas";
$subseccion3_4 = "Otra formación universitaria de posgrado";
$subseccion3_5 = "Formación especializada, continuada, técnica, profesionalizada, de reciclaje y actualización (distinta a la formación académica)";
$subseccion3_6 = "Formación sanitaria especializada";
$subseccion3_7 = "Formación sanitaria en I+D";
$subseccion3_8 = "Cursos y seminarios recibidos de perfeccionamiento, innovación y mejora docente, nuevas tecnologías, etc., cuyo objetivo sea la";
$subseccion4_1 = "Dirección de tesis doctorales y/o proyectos fin de carrera";
$subseccion4_2 = "Formación académica impartida";
$subseccion4_3 = "Formación sanitaria especializada impartida";
$subseccion4_4 = "Formación sanitaria en I+D, y/o posformación sanitaria especializada en I+D impartida";
$subseccion4_5 = "Tutorías académicas de estudiantes";
$subseccion4_6 = "Cursos y seminarios impartidos orientados a la formación docente universitaria";
$subseccion4_7 = "Publicaciones docentes o de carácter pedagógico, libros, artículos, etc.";
$subseccion4_8 = "Participación en proyectos de innovación docente";
$subseccion4_9 = "Participación en congresos con ponencias orientadas a la formación docente";
$subseccion4_10 = "Premios de innovación docente recibidos";
$subseccion4_11 = "Otras actividades/meritos no incluidos en la relación anterior";
$subseccion4_12 = "Aportaciones más relevantes de su CV de docencia";
$subseccion5_1 = "Proyectos de I+D+i financiados en convocatorias competitivas de Administraciones o entidades públicas y privadas";
$subseccion5_2 = "Contratos, convenios o proyectos de I+D+i no competitivos con Administraciones o entidades públicas o privadas";
$subseccion5_3 = "Propiedad industrial e intelectual";
$subseccion5_4 = "Grupos/equipos de investigación, desarrollo o innovación";
$subseccion5_5 = "Obras artísticas dirigidas";
$subseccion5_6 = "Resultados tecnológicos derivados de actividades especializadas y de transferencia no incluidos en apartados anteriores";
$subseccion6_1 = "Indicadores generales de calidad de la producción científica";
$subseccion6_2 = "Publicaciones, documentos científicos y técnicos";

```

Figura 6.39: Variables con todas las secciones y subsecciones encontradas en el editor de Fecyt

Por otro lado, es necesario obtener el título del elemento que haga referencia a un código en concreto. Por lo general, el título se encuentra en el nodo *Title*, dentro de otro llamado *Name* y finalmente en otro conocido como *Item*. Esto puede cambiar dependiendo del código que se analice, por lo que es necesario indicar para cada código como encontrar el título del elemento, con el objetivo de mostrarlo al usuario antes de escoger a donde pertenecerá el elemento.

Finalmente, el vector de *categorias* está pensando para indicar las posibles categorías a las que puede pertenecer un determinado elemento. Si un elemento pertenece a dos categorías diferentes el usuario deberá resolver el conflicto escogiendo una de las dos categorías, tal y como se explicó en la sección del lado del cliente.

En la figura 6.40 se ve un ejemplo de los campos comentados hasta ahora para el código 010.010.000.000. Este código solo puede pertenecer a una categoría, pero podría darse el caso de que perteneciese a más de una de las listadas en 6.2.2.

```
//010.010.000.000: Otros -> Situación profesional -> Situación profesional
"010.010.000.000" => [
  "seccion" => $seccion2,
  "subseccion" => $subseccion2_1,
  "titulo" => ["Title", "Name", "Item"],
  "categorias" => [
    "1" => [...],
  ],
],
```

Figura 6.40: Índices y valores para el índice del código de elemento

Como se ha comentado, el índice de categoría tiene como valor un vector asociativo, en el que se muestran los siguientes campos:

- **nombre:** El nombre de la categoría de SICA a la que pertenece el código (Otros, Publicación en revista, etc.).
- **tipos_item_sica:** Muestra todos los posibles tipos de elementos a los que puede pertenecer el código del elemento CVN en SICA. A su vez, mostrará todos los posibles subtipos de elemento a los que puede pertenecer el elemento del CVN según el tipo escogido.
- **campos_mostrar:** Contiene todos los posibles campos a mostrar para una categoría, tipo de elemento y subtipo de elemento, indicando el tipo de campo HTML y los posibles posibles que pueden encontrarse en el CVN.
- **pasos:** Muestra todas las peticiones a través de Internet que hay que realizar para poder importar correctamente un elemento a la plataforma SICA.

En el diagrama de la figura 6.41 se muestra la estructura para el vector de *categorías*:

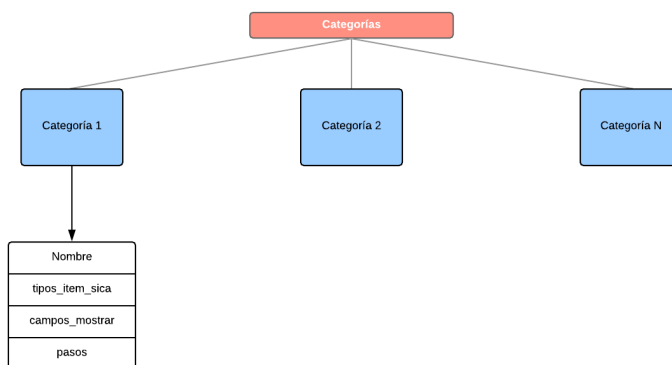


Figura 6.41: Diagrama en el que se muestra el vector de *categorías*

En la figura 6.42 se muestra un ejemplo de los índices que son mostrados para el índice de *categorías* visto anteriormente.

```

//010.010.000.000: Otros -> Situación profesional -> Situación profesional
"010.010.000.000" => [
  "seccion" => $seccion2,
  "subseccion" => $subseccion2_1,
  "titulo" => ["Title", "Name", "Item"],
  "categorias" => [
    "1" => [
      "nombre" => "Otros",
      "tipos_item_sica" => [...],
      "campos_mostrar" => [...],
      "pasos" => [...],
    ],
  ],
],
1,
1,
1,

```

Figura 6.42: Índices y valores para el índice de *categorías*

El índice *tipos_item_sica* contiene los posibles tipos de *item* que pertenecen a un determinado código para la categoría escogida. Además, por lo general, un tipo de elemento en SICA contiene a su vez diferentes subtipos en los que se puede clasificar el elemento curricular que se desee subir a la plataforma. Estos tipos y subtipos pueden observarse en la plataforma SICA, como se pudo observar en la figura 6.11, dónde cada elemento del seleccionable (*select*) tiene un valor en forma de identificador numérico que será necesario proporcionar a SICA a la hora de realizar las peticiones de

subida de un *item*.

En el diagrama 6.43 se muestra la estructura para el vector *tipos_item_sica*, en el que se visualiza como puede haber un número N de tipos de *item* que a su vez pueden tener otro número N de subtipos de *item*.

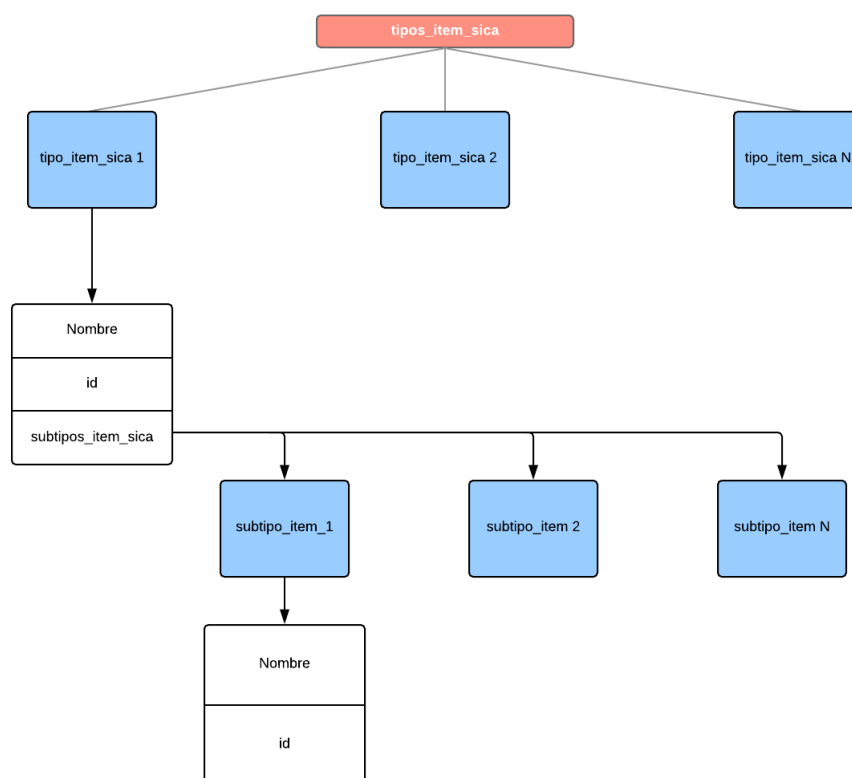


Figura 6.43: Diagrama en el que se muestra el vector de *tipos_item_sica*

En la figura 6.44 se observa como un determinado código puede pertenecer a dos tipos de *item* diferentes, y además cada tipo contiene diferentes subtipos. Tanto tipos como subtipos de *item* tienen su propio identificador, que en realidad es un valor del campo *tipo select* que aparece en la página web de SICA a la hora de seleccionar elemento como se veía en la figura 6.11, y que puede ser obtenido inspeccionando dicho campo, como se observa en la figura 6.45.

```

"categorias" => [
  "1" => [
    "nombre" => "Otros",
    "tipos_item_sica" => [
      "1" => [
        "nombre" => "Formación Universitaria",
        "id" => 19,
        "subtipos_item_sica" => [
          "1" => [
            "nombre" => "Titulado medio",
            "id" => 46,
          ],
          "2" => [
            "nombre" => "Titulado superior",
            "id" => 47,
          ],
          "3" => [
            "nombre" => "Otra titulación",
            "id" => 44,
          ],
        ],
      ],
      "2" => [
        "nombre" => "Idiomas",
        "id" => 23,
        "subtipos_item_sica" => [
          "1" => [
            "nombre" => "Conocimiento de idiomas",
            "id" => 41,
          ],
        ],
      ],
    ],
  ],
],

```

Figura 6.44: Índices y valores para el índice de *tipos_item_sica*

```

<option value="org.jboss.seam.ui.NoSelectionConverter.noSelectionValue" selected=
"selected">Selecione un tipo de ítem...</option>
<option value="1">Actividades genéricas docentes</option>
<option value="2">Actividades genéricas sanitarias</option>
<option value="3">Becas y Reconocimientos</option>
<option value="4">Cargos y Actividades Profesional Anteriores</option>
<option value="5">Catálogo de Obra Artística</option>
<option value="6">Colaboraciones con Centros I+D</option>
<option value="7">Comités científicos de revistas</option>
<option value="8">Comités, Foros, Redes y Sociedades</option>
<option value="9">Conferencias Impartidas</option>
<option value="10">Congresos, Exposiciones, etc.</option>
<option value="11">Convenios y Contratos</option>
<option value="12">Docencia impartida</option>
<option value="13">Edición de Actas de Congresos, etc.</option>
<option value="14">Edición de Revistas</option>
<option value="15">Ensayos Clínicos</option>
<option value="16">Entrevistas en Medios de Comunicación</option>
<option value="17">Estancias</option>
<option value="18">Experiencia en Evaluación</option>
<option value="19">Formación Universitaria</option> == %0
<option value="20">Gestión en I+D+I</option>
<option value="21">Grupos, Equipos, etc.</option>
<option value="22">Guías, protocolos, etc. sanitarias</option>
<option value="23">Idiomas</option>

```

Figura 6.45: Identificadores de tipo *item* en SICA

En la pantalla de pre-selección que se comentó en la sección del lado del cliente, se muestran las posibles categorías, tipos de *item* y subtipos de *item* de cada elemento en función de su código basándose en el vector de *tipos_item_sica* que se acaba de explicar.

Por otro lado, el índice de *campos_mostrar* mostrará todos los posibles campos a mostrar para una determinada categoría, tipo *item* o subtipo *item*. Estos campos son los que aparecen en SICA cuando escoges en donde clasificar el elemento que se desea subir, visualizados en la figura 6.12. Como se observa, cada campo puede ser un *input* de tipo texto (*text*), fecha (*date*), seleccionable (*select*), cuadro de texto (*textarea*) o un elemento oculto (*hidden*).

En la figura 6.46 se muestra la estructura general para el índice de *campos_mostrar*, en el cual se observan los atributos principales que podría tener cada campo.

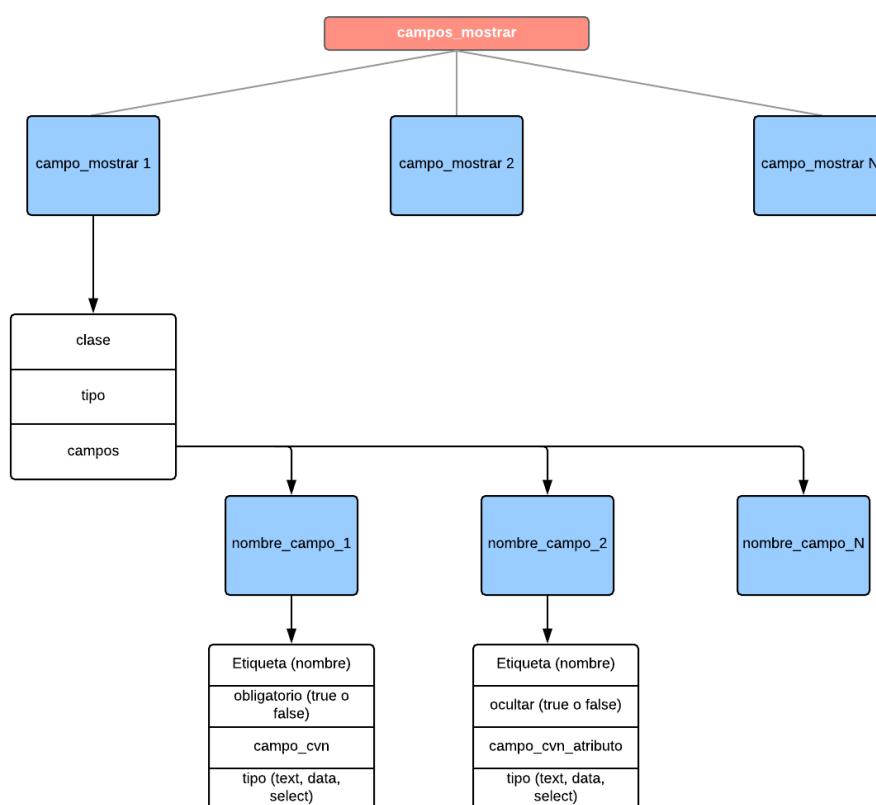


Figura 6.46: Diagrama en el que se muestra el vector de *campos_mostrar*

Cabe destacar que existen más atributos para cada campo que no se han

mencionado, ya que tienen una elevada complejidad. Esto se debe a que el objetivo es mostrar la mayor cantidad de información posible y cada campo puede tener unos atributos diferentes a otro, según como se muestre en SICA.

Por lo tanto, para representar todos estos campos en el código HTML que visualizará el usuario, se ha creado una clase llamada *HTMLCreator.php* que contiene las funciones necesarias para imprimir por pantalla cada campo que pueda existir en SICA, incluyendo diversas propiedades, como la clase, id, o el valor obtenido del CVN, además de indicar si dicho campo es obligatorio o no.

En la figura 6.47 se muestra un ejemplo de cómo se han añadido los campos a mostrar para un determinado elemento.

```
"campos_mostrar" => [
  "1" => [...],
  "2" => [
    "clase" => "segunda_pagina",
    "tipo" => "div",
    "campos" => [
      "formFatherNewItem:calendarLB22InputDate" => [
        "nombre" => "Fecha inicio",
        "obligatorio" => 1,
        "campo_cvn" => ["Date", "StartDate", "DayMonthYear", "Item"],
        "tipo" => "date",
      ],
      "formFatherNewItem:calendarLB22InputCurrentDate" => [
        "nombre" => "Fecha inicio actual",
        "obligatorio" => 1,
        "ocultar" => 1,
        "valor" => date("m/Y"),
      ],
      "formFatherNewItem:textPlaceLB24" => [
        "nombre" => "Lugar",
        "obligatorio" => 0,
        "campo_cvn" => ["Place", "City", "Item"],
      ],
      "formFatherNewItem:aneparea" => [
        "nombre" => "Áreas ANEP",
        "tipo" => "select",
        "obligatorio" => 0,
        "opciones" => [
          "Seleccione un área..." => "",
          "Área de Agricultura" => 1,
          "Área de Biología Molecular, Celular y Genética" => 2,
          "Área de Biología Vegetal, Animal y Ecología" => 3,
          "Área de Ciencia y Tecnología de Alimentos" => 4,
          "Área de Ciencia y Tecnología de Materiales" => 5,
          "Área de Ciencias de la Computación y Tecnología Informática" => 6,
          "Área de Ciencias de la Tierra" => 7,
        ],
      ],
    ],
  ],
]
```

Figura 6.47: Índices y valores para el índice de *campos_mostrar*

Como se puede observar, se indica el nombre de cada campo, el título para cada uno, si es obligatorio o no, si debe permanecer oculto en el HTML, el valor encontrado en el CVN (*campo_cvn*), o el valor por defecto que tendrá algún campo (*valor*). Además, se separa cada conjunto de campos a mostrar como un vector diferente, enumerados desde 1, cada cual con su propio contenedor (*div*), clase o título.

Los casos más complicados en los campos a mostrar son los seleccionables, ya que incluyen diferentes opciones de las que habrá que obtener su valor. Un ejemplo de esto es el seleccionable de áreas ANEP, que aparece en todos los elementos de SICA y que puede cambiar el valor de sus opciones

en cada elemento, por lo que habrá que dedicar tiempo a obtener las posibles opciones y sus valores.

Un caso especial es el de los autores en SICA, ya que existen diferentes elementos que pueden tener uno o más autores, información que puede verse reflejada en el CVN. En ese caso, es necesario recorrer el CVN-XML en busca de todos los posibles autores incluidos, mostrando al usuario utilizando un bucle dicha información.

Finalmente, se explicarán los campos contenidos en el índice de *pasos*. Esto conlleva la mayor complicación de la herramienta, ya que se indican los pasos a seguir para poder subir un elemento a SICA correctamente. Completarlo requiere del análisis del tráfico explicado en 6.2.2, ya que cada elemento tiene sus particularidades para ser subido.

En primer lugar, se ha creado un vector por cada paso que haya que realizar para subir un elemento, donde cada uno es una petición POST que hace uso de la subida de un formulario o una llamada a través de AJAX con JavaScript. Cada paso tiene una descripción para conocer lo que se realiza en ese momento en caso de que requiera una revisión o modificación. Además, se incluye la URL a visitar en cada paso, ya que ésta puede variar según la página de SICA que se esté visitando.

Además, en cada paso se incluyen los campos que deben ser incluidos en la petición que se realiza, siendo algunos de ellos los mostrados en el HTML del usuario y otros que no aparecen pero que deben ser incluidos, motivo por el cual es necesario el análisis del tráfico.

En la figura 6.48 se muestra la estructura del vector de *pasos*, indicando los atributos más importantes, como la URL a la que se envía la petición, los campos que se han mostrado por pantalla al usuario (*campo_mostrados*) y campos que es necesario enviar para cada paso.

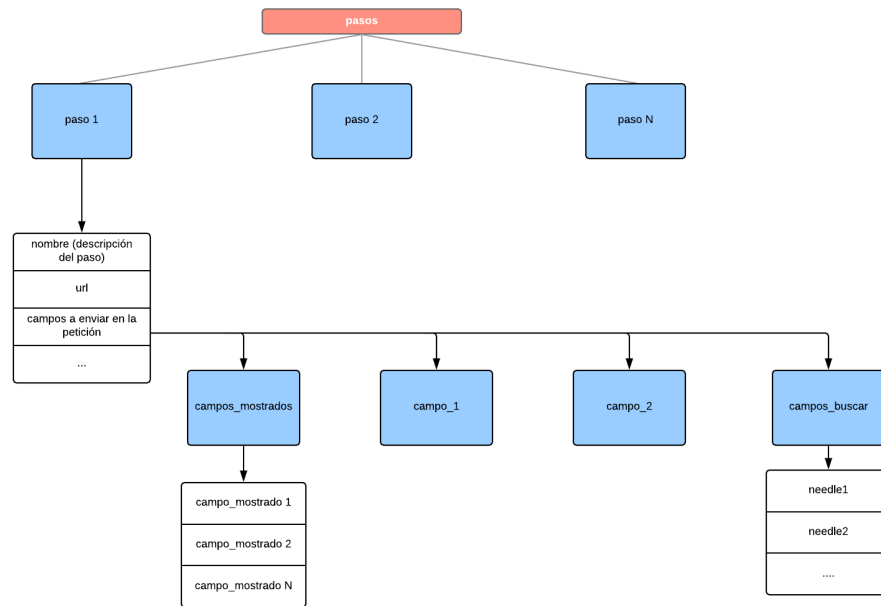


Figura 6.48: Diagrama en el que se muestra el vector de *pasos*

Es importante destacar que, como se comentó en la sección del análisis del tráfico, existen campos cuyo nombre es dinámico y podría cambiar con el paso del tiempo, motivo por el cual es necesario encontrarlos antes de realizar la petición. Esto se realizará con expresiones regulares, indicando las cadenas de texto entre las que se encuentra el nombre del campo a buscar.

Por último, cada conjunto de pasos varía según el tipo o subtipo de elemento, por lo que en ocasiones es necesario indicar para que tipo o subtipo están indicados los pasos de subida de elementos.


```

"pasos" => [
  "1" => [
    "nombre" => "Entrar a Otros",
    "url" => "https://sica2.cica.es/investigacion/pages/workforms/newsWorkforms/newItem.jsf",
    "campos" => null,
  ],
  "2" => [
    "nombre" => "Primer AJAX. form:workformtype",
    "url" => "https://sica2.cica.es/investigacion/pages/workforms/newsWorkforms/newItem.jsf",
    "viewstate" => 1,
    "campos" => [
      "mostrados" => ["form:workformtype"],
      "autoScroll" => null,
      "form" => "form",
      "AJAXREQUEST" => "_viewRoot",
      "ajaxSingle" => "form:workformtype",
      "campos_buscar" => [
        "0" => [
          "needle1" => \"similarityGroupId\": \"form:\",
          "needle2" => \"\",
          "key_inicial" => "form:",
          "valor_inicial" => "form:",
          "key" => true,
          "valor" => true,
          "html_buscar" => 0,
          "alternativa" => "j_id92",
        ],
      ],
    ],
  ],
],
1,

```

Figura 6.49: Índices y valores para el índice de *pasos*

En la figura 6.49 se muestra un ejemplo de dos pasos a seguir para poder subir un elemento. Como se observa aparecen todos los campos que es necesario incluir para realizar la petición a la URL indicada.

El índice de *campos_buscar* contiene las indicaciones para encontrar el nombre de un campo, que será necesario para incluir una clave y un valor en la petición. Para ello, se declara el inicio de la cadena a buscar (*needle1*) y el final de la cadena (*needle2*), además de indicarse la clave y el valor inicial para dicho parámetro a incluir. Por lo general, estos parámetros comienzan con la palabra *form:* seguido de la cadena buscada, aunque no siempre es así. Por lo tanto, con los índices *key* y *value* se indica si se desea incluir la cadena en los parámetros de la petición. Además, todos los códigos HTML que se generan en cada paso son guardados en un vector, por lo que es importante indicar cual de los códigos HTML es en el que hay que buscar dicho campo, ya que conforme se envían peticiones al servidor se reciben respuestas con diferente código HTML. Por último, en caso de que SICA cambie lo suficiente su código HTML como para no encontrar ninguna cadena con los métodos propuestos, se incluye una alternativa con el valor actual que tiene en su página.

Una vez explicado cómo se almacena la información de todos los posibles elementos importables a SICA se detallarán las otras partes del servidor, consistentes en el análisis del CVN introducido y su importación a SICA.

En primer lugar, se deberá iniciar sesión en la herramienta con las credenciales de SICA, tal y como explicaba en la sección del lado del cliente. Este inicio de sesión queda registrado en unas variables de PHP conocidas como *\$_SESSION* [31], una variable global que almacena datos para usuarios

de manera individual usando un identificador de sesión único. Esto puede utilizarse para hacer persistente la información de estado entre peticiones de páginas. En esta variable se guardará la información de usuario y contraseña del usuario, que servirá para poder realizar la importación posteriormente. Esta variable de sesión perdura hasta que se cierra el navegador o cuando PHP termina de ejecutar el script, aunque es posible cerrarla manualmente, aunque no este último caso no se ha considerado para permitir al usuario importar su CVN de una manera más rápida, sin necesidad de iniciar sesión continuamente.

Para cerrar sesión el usuario deberá pinchar sobre el botón de 'Cerrar sesión' que aparecía en la figura 6.23, momento en el cual se enviará al usuario a una página PHP que se encargará de vaciar las variables de sesión que guardaban la información referente al nombre de usuario y la contraseña.

Por otro lado, una vez que se ha iniciado sesión, al introducir un fichero CVN-PDF o CVN-XML en la herramienta, se envían los datos a una página PHP en la cual se comprueba si la extensión del archivo es PDF o XML. En caso de que la extensión sea PDF se busca el contenido del código codificado en Base64 que incluye el archivo comprimido y que a su vez contiene el archivo CVN-XML. Si se encuentra el CVN-XML se procede a procesar todo el contenido del CVN, incluyendo el nodo de *Version*, *Agent* y todos los *CvnItem*. La información de los nodos *Version* y *Agent* está recogida en el servidor, aunque no se utiliza en ningún momento debido a su poca utilidad para la herramienta. Los únicos nodos que se procesan actualmente son todos los *CvnItem*, incluyendo el nodo XML completo en forma de vector.

Una vez analizados todos los nodos de *CvnItem* se comprueba si alguno de ellos presenta conflictos como se explicó en 6.2.3 y si no, se pasa a seleccionar los elementos que se importarán. En caso de que haya conflictos, se muestran únicamente aquellos elementos que pueden pertenecer a más de una categoría, tipo de elemento o subtipo de elemento en SICA.

En la página de selección de elementos se mostrará un listado que se basará en los índices de *campos_mostrar* que se explicó anteriormente, identificando el tipo de entrada HTML, si su valor se encuentra en el CVN y analizando todas las posibles opciones que se encuentran dentro del vector de variables. Una vez analizados todos los índices del vector *campos_mostrar* del *item* se mostrarán por pantalla al usuario con su valor correspondiente en caso de que exista. La búsqueda de valores se realizará indicando los nodos en los que hay buscar o por un código de atributo que identifica un único nodo. Además, en esta página se realizan otro tipo de procesamientos complejos como es el caso de la búsqueda de autores en el CVN que requieren mostrar, a través de un bucle, todos los valores encontrados.

Por último, los elementos seleccionados pasarán a la página *procesa_importar.php* que procesará todos los datos enviados haciendo uso del índice de *pasos* del vector de variables comentado anteriormente.

Al completar todas las peticiones de todos los elementos que se desean

enviar la herramienta habrá cumplido su función y se habrán importado todos los elementos escogidos del CVN a SICA. En caso de que algún elemento no pueda añadirse, se guardará en un vector de fallos, indicando su título. Cuando termine la importación si el vector de fallos contiene algún elemento se mostrará al usuario por pantalla cuales son aquellos que no pudieron ser encontrados. Estos fallos se encontrarán buscando en el código HTML, ya que cuando un elemento da un error al importarse a SICA aparece un mensaje de error en la página que puede ser buscado a través de expresiones regulares.

Capítulo 7

Fase de pruebas.

Este capítulo describe el proceso que se ha llevado a cabo para realizar la fase de pruebas del proyecto. Esta fase de pruebas tiene como objetivo comprobar el comportamiento del sistema completo y solucionar posibles errores.

A lo largo de la implementación del proyecto, se han ido realizando distintas pruebas, con el fin de comprobar que lo que se programaba funcionaba correctamente y sin fallos. Además, se corregían distintos aspectos de implementación de cara al uso de la aplicación por parte de un usuario.

Las pruebas se han realizado utilizando XAMPP, un servidor independiente de plataforma de software libre que se ha utilizado para emular el servidor en una página web y hacer distintas pruebas. Con esto, se puede emular un servidor de manera local que servirá para comprobar el funcionamiento de la herramienta.

A continuación, se muestran las diferentes pruebas que se han llevado a cabo a lo largo del proyecto:

Prueba 1: Análisis de tráfico en SICA

En primer lugar, es importante destacar las pruebas realizadas con mitmproxy para analizar el tráfico en SICA. Para ello, es necesario configurar la red área local (LAN) del dispositivo para que utilice un servidor proxy, indicando una dirección y puerto. La dirección será 127.0.0.1, que hace referencia a *localhost*, es decir, la dirección que apunta al ordenador. Por otro lado, el puerto será el escogido con mitmproxy a la hora de lanzarlo. Puede verse la configuración de red de área local en la figura 7.1

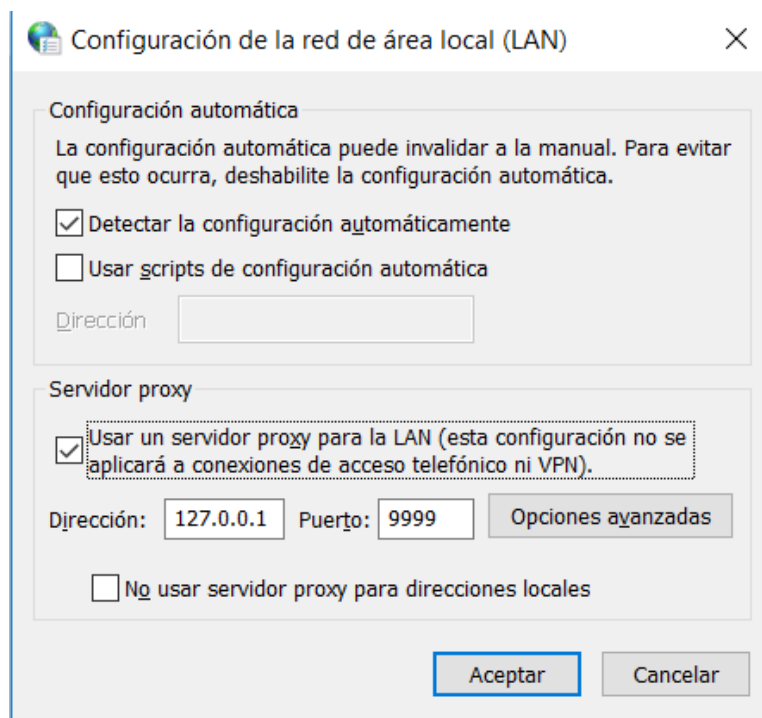


Figura 7.1: Configuración de área local

Seguidamente, es necesario lanzar mitmproxy en su versión web con el siguiente comando a través del CMD, donde el parámetro `-p` hace referencia al puerto. Para ello se requiere que mitmproxy esté instalado en el equipo.

```
mitmweb -p 9999
```

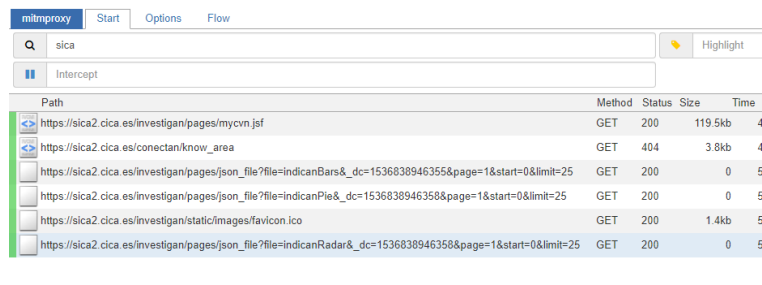
Este comando abrirá en el navegador por defecto del equipo la interfaz gráfica que permitirá observar todas las peticiones y respuestas enviadas a través del navegador. Con esto, se realizaron todas las pruebas relacionadas con el análisis del tráfico, que han sido útiles y necesarias para el desarrollo de la herramienta.

Prueba 2: Importación de elementos encontrados en el CVN a SICA

Una vez explicado cómo se utilizó mitmproxy, se procede a explicar la manera en la que se realizó el guardado de todos los elementos importables a SICA en la herramienta, conociendo todos los campos a mostrar al usuario, su valor en el CVN y las peticiones que eran necesarias realizar.

El procedimiento ha sido el siguiente:

- Iniciar mitmproxy y filtrar por SICA

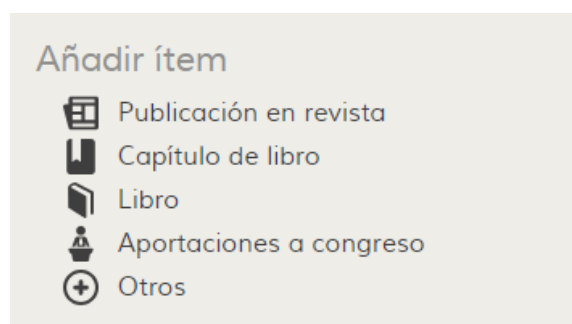


The screenshot shows the mitmproxy web interface. At the top, there's a search bar with 'sica' entered and a 'Highlight' button. Below it is an 'Intercept' button. The main part of the interface is a table of intercepted requests.

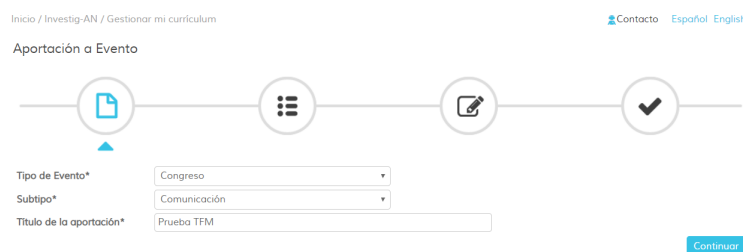
Path	Method	Status	Size	Time
https://sica2.cica.es/investigan/pages/mycvn.jsf	GET	200	119.5kb	4s
https://sica2.cica.es/conectan/know_area	GET	404	3.8kb	4s
https://sica2.cica.es/investigan/pages/json_file?file=indicanBars&_dc=1536838946355&page=1&start=0&limit=25	GET	200	0	5s
https://sica2.cica.es/investigan/pages/json_file?file=indicanPie&_dc=1536838946358&page=1&start=0&limit=25	GET	200	0	5s
https://sica2.cica.es/investigan/static/images/favicon.ico	GET	200	1.4kb	5s
https://sica2.cica.es/investigan/pages/json_file?file=indicanRadar&_dc=1536838946358&page=1&start=0&limit=25	GET	200	0	5s

Figura 7.2: Mitmweb con filtrado por SICA

- Acceder a SICA y escoger la categoría del *item* a subir

Figura 7.3: Categoría del *item* a escoger en SICA

Además, habrá que escoger el tipo de *item* y el subtipo de *item*, si la categoría lo permite:



The screenshot shows a form titled 'Aportación a Evento'. It has a progress bar at the top with four steps: 1. Tipo de Evento (selected), 2. Subtipo, 3. Título de la aportación, and 4. Continuar. Below the progress bar, there are three input fields:

- Tipo de Evento*: Congreso (dropdown menu)
- Subtipo*: Comunicación (dropdown menu)
- Título de la aportación*: Prueba TFM (text input field)

At the bottom right, there is a blue button labeled 'Continuar'.

Figura 7.4: Tipo y subtipo del *item* a escoger en SICA

- Rellenar todos los campos posibles del elemento en SICA

PROYECTO FIN DE CARRERA

Área de tipología | **Área de título y descripción** | Rol en el ítem curricular | Indicadores de calidad

Título *

Idioma

Fecha de lectura *

Centro de lectura *

Temática

Resumen

Lugar

Figura 7.5: Campos rellenados para un *ítem* de SICA

- Observar el nombre de cada uno de los campos que han sido rellenados

Por ejemplo, para el campo 'Resumen' de la figura 7.5 vista anteriormente el campo tiene como nombre *formFatherNewItem:descriptionLB27*.

```
<textarea id="formFatherNewItem:
descriptionLB27" name="formFatherNewItem:
descriptionLB27" class="ancho95p
areaWithoutWidth areaNoResize" rows="10">
</textarea>
```

Figura 7.6: Nombre de un campo en SICA

- Insertar en la herramienta el nombre de los campos y sus posibles valores (en caso de los *selects*) en el índice de *campos_mostrar*. Siguiendo con el ejemplo del campo 'Resumen', en la figura 7.7 se observa el campo añadido y cómo obtener su valor del CVN a partir del atributo *campo_cvn*.

```
"campos_mostrar" => [
  "1" => [...],
  "2" => [
    "clase" => "segunda_pagina",
    "tipo" => "div",
    "campos" => [
      "formFatherNewItem:descriptionLB27" => [
        "nombre" => "Resumen",
        "tipo" => "textarea",
        "campo_cvn" => ["Description", "Item"],
      ],
    ],
  ],
]
```

Figura 7.7: Campo añadido al vector *campos_mostrar*

- Guardar el elemento en SICA

GESTIONAR MI CURRÍCULUM

Use el buscador para buscar el ítem curricular que desee editar, consultar o borrar. Además, dispone de una serie de accesos directos para poder moverse a otros apartados de la aplicación con total comodidad.

Estado	Tipo	Título	Rol	Fecha	Detalle
●	Proyecto fin de carrera	Herramienta de importación de CVN a SICA	Director	14/09/2018	Q ✎ 🗑

Figura 7.8: *Item* añadido a SICA

- Insertar en la herramienta todas las peticiones y parámetros que son necesarios enviar para subir el elemento. Esto se realiza en el vector de *pasos*.

```
"pasos" => [
  "1" => [...],
  "2" => [...],
  "3" => [...],
  "4" => [...],
  "5" => [...],
  "6" => [...],
  "7" => [...],
  "8" => [...],
  "9" => [...],
  "10" => [
    "nombre" => "Insercción de ITEM.",
    "url" => "https://sica2.cica.es/investigacion/pages/workforms/newsWorkforms/fatherNewItem.jsf",
    "campos" => [
      "mostrados" => [
        "formFatherNewItem:aneparea",
        "formFatherNewItem:filterValueLB22",
        "formFatherNewItem:externalLB23",
        "formFatherNewItem:calendarLB24InputDate",
        "formFatherNewItem:calendarLB24InputCurrentDate",
        "formFatherNewItem:calendarLB25InputDate",
        "formFatherNewItem:calendarLB25InputCurrentDate",
        "formFatherNewItem:entityLB32",
        "formFatherNewItem:descriptionLB27",
        "formFatherNewItem:ecdimensionLB41",
        "formFatherNewItem:j_id327",
        "formFatherNewItem:j_id335",
      ],
      "mostrados_buscar" => [
        "0" => [
          "campo_mostrado" => "formFatherNewItem:j_id327",
          "needle1" => '<select name="formFatherNewItem:',
          "needle2" => '"',
          "key_inicial" => "formFatherNewItem:",
          "key" => true,
        ]
      ]
    ]
  ]
]
```

Figura 7.9: Pasos añadidos a la herramienta para importar un *item* a SICA

- Exportar el CVN desde SICA con ese elemento en formato CVN-XML

GENERAR MI CURRÍCULUM

Seleccione los elementos a exportar de las opciones disponibles a la derecha de la pantalla. Además, puede usar el buscador básico o avanzado para encontrar sus ítems curriculares y así decidir cuáles exportar. Por último, seleccione el formato de exportación que desee y pulse el botón exportar.

Tipo	Título	Rol	Fecha	
Proyecto fin de carrera	Herramienta de importación de ...	Director	14/09/2018	<input checked="" type="checkbox"/>

Criterios de exportación

☒ Todas las fechas
 ☐ Intervalo personalizado

Seleccione un formato de exportación: CVN-XML

Figura 7.10: Exportación de CVN-XML desde SICA

- Observar todos los valores insertados en SICA que se han extrapolado al CVN-XML

```

<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<ns2:CVN xmlns:ns2="http://cv.normalizado.org/cvn">
  <Version...>
  <Agent...>
  <CvnItem...>
  <CvnItem>
    <CvnItemID>
      <CVNPK code="030.040.000.000">
        <Item>030.040.000.000</Item>
      </CVNPK>
      <CodeCvnItem code="030.040.000.000">
        <Item>010</Item>
      </CodeCvnItem>
      <ExperienceType code="030.040.000.000">
        <Item>000</Item>
      </ExperienceType>
      <ResultType code="030.040.000.000">
        <Item>010</Item>
      </ResultType>
    </CvnItemID>
    <Subtype>
      <SubType1 code="030.040.000.010">
        <Item>084</Item>
      </SubType1>
    </Subtype>
    <Title>
      <Name code="030.040.000.030">
        <Item>Herramienta de importación de CVN a SICA</Item>
      </Name>
      <Type>000</Type>
    </Title>
    <Entity>
      <EntityName code="030.040.000.080" obligatory="true">
        <Item>ETSIIT</Item>
      </EntityName>
    </Entity>
  </CvnItem>
</ns2:CVN>

```

Figura 7.11: Explotación de CVN-XML del *item* añadido a SICA

- Añadir en la herramienta el identificador para ese *item* y los valores que se encuentran en el CVN-XML para los campos añadidos anteriormente

```
//030.040.000.000: Tesis, Tesinas, DEA, etc. -> Proyecto de fin de carrera. Tesina. Tesis Doctoral.
"030.040.000.000" => [
  "seccion" => $seccion4,
  "subseccion" => $subseccion4_1,
  "titulo" => ["Title", "Name", "Item"],
  "incluye_subtipos_subtipo" => ["Subtype", "SubType1", "Item"],
  "categorias" => [
    "1" => [
      "nombre" => "Otros",
      "tipos_item_sica" => [...],
      "campos_mostrar" => [...],
      "pasos" => [...],
    ],
  ],
],
```


Figura 7.12: Identificación de *item* y los valores de los campos

- Insertar el CVN-XML exportado desde SICA en la herramienta para realizar la importación



Figura 7.13: CVN-XML añadido en la herramienta

- Rellenar todos los campos del *item* y pulsar sobre 'Importar' en la herramienta

 **Herramienta de importación de cvn a sica** (Actividad docente - Dirección de tesis doctorales y/o proyectos fin de carrera)

Tipo de Ítem (*)
Tesis, Tesinas, DEA, etc. ▼

Subtipo (*)
Proyecto fin de carrera ▼

Título del ITEM (*)
Herramienta de importación de cvn a sica

Áreas ANEP
Seleccione un área... ▼

Idioma
Seleccione un idioma... ▼

Fecha de lectura (*)
dd/mm/aaaa

Centro de lectura
Etsiit

Temática

Resumen

Lugar

Figura 7.14: Selección de *item* añadido a la herramienta

- Esperar a que el *item* termine de importarse, evento que ocurrirá una vez desaparezca la imagen del *loader* de la figura 7.15

Temática
Temática
Resumen
Resumen
Lugar
Granada
Calificación
9
Mención Internacional
Sí
Doctorado Europeo
Sí
Rol del Autor del CVN
Rol autor (*)
Autor/a
Tipo de dedicación autor
Tiempo completo

Importar a SICA




Figura 7.15: Momento de importación en la herramienta

- Comprobar si el elemento del CVN-PDF se ha importado en SICA correctamente

Importar a SICA

Todos los items han sido importados correctamente

Figura 7.16: Importación correcta del *item* seleccionado

Además, se comprueba si en SICA aparece el elemento y con

todos los campos que se habían rellenado:

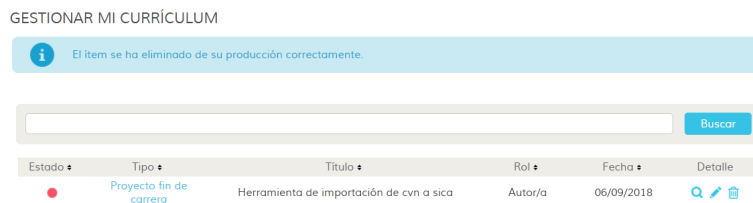


Figura 7.17: *Item* importado a SICA correctamente

Además, han sido necesarias pruebas para comprobar que se podían importar dos o más elementos incluidos en un CVN. Estas pruebas se realizaron una vez que todos los elementos importables a SICA habían sido comprobados y se concluyó que la importación funcionaba correctamente. Para estas pruebas únicamente se añadían a SICA diferentes elementos, siempre más de uno, y seguidamente se exportaba el contenido del currículum al formato CVN-PDF. Con el CVN-PDF generado, se añadía a la herramienta y se comprobaba que se clasificaban los elementos correctamente, además de comprobar si se detectaban conflictos en caso de que un elemento pudiese pertenecer a diferentes tipos en SICA. Una vez que se tenían todos los elementos cargados en la herramienta se seleccionaban dos o más y se procedía a la importación, rellenando previamente todos los campos de carácter obligatorio. Estas pruebas sirvieron también para comprobar que ocurriría en caso de que fallara la importación de un determinado elemento, por cualquier tipo de error. Se concluyó que en caso de que la importación de un elemento fallase se pasaría al siguiente elemento a importar, sin que el fallo afectase al resto.

Prueba 3: Diseño de la herramienta

Una vez que la herramienta se había implementado correctamente, se realizaron numerosas pruebas con el diseño de la herramienta, buscando diferentes plantillas acordes a la naturaleza de la herramienta. Se requirió de diversas pruebas para definir el diseño al actual, además de los ajustes de CSS para ajustar todos los elementos HTML acordes a todos los tipos de pantallas haciendo uso de Bootstrap. Aunque el diseño no es el punto principal de la herramienta, se ha intentado que tenga un diseño agradable y sencillo. Todo el diseño puede ser observado en la sección 6.2.3 del capítulo de implementación.

Capítulo 8

Conclusiones y líneas futuras.

En este capítulo se desarrollarán las conclusiones sacadas de este proyecto y las posibles líneas futuras que pueden ser requeridas o implementadas.

8.1. Conclusiones

En el presente Trabajo de Fin de Máster se ha desarrollado una herramienta con la capacidad de analizar un fichero en formato CVN-PDF e importar a SICA la información obtenida del mismo.

El objetivo propuesto ha sido desarrollo con éxito, elaborándose una herramienta utilizada a través de una página web que actuará como analizador de ficheros CVN escogidos por el usuario y permitirá importarlos a SICA, mostrando todos los campos necesarios. Esto ha requerido realizar una parte dedicada al cliente, utilizando tecnología HTML, JavaScript y CSS, y una parte servidora en donde se realiza todo el procesamiento, como es el caso del análisis del CVN y la importación a SICA, utilizando para ello el lenguaje PHP.

Las contribuciones más destacables de este proyecto son las siguientes:

- Solventa el problema de no poder importar ningún tipo de archivo en la plataforma SICA, siendo posible ahora importar un fichero CVN correctamente formado.
- Permite analizar todos los elementos incluidos en un CVN, categorizando la información en categorías, tipos de *item* y subtipos de *item* acorde a la estructura de SICA. Se intenta obtener el máximo posible de información incluida en el CVN, aunque como se ha explicado a lo largo del proyecto, existen numerosos campos en SICA que no pueden ser extrapolados al CVN, por lo que dicha información deberá ser rellenada por el usuario.

- Se ha realizado una investigación exhaustiva sobre el análisis del tráfico de todos los posibles elementos que pueden ser añadidos a SICA, explicando el proceso a lo largo de este proyecto. Esto supone que cualquiera que quiera dar continuidad al proyecto o aplicarlo a otras plataformas es capaz de hacerlo siguiendo los pasos dados.
- La herramienta desarrollada es intuitiva para el usuario, teniendo una interfaz visualmente sencilla con pocas opciones. Únicamente se permite iniciar sesión con las credenciales de SICA, seleccionar el fichero CVN, clasificar los elementos en su tipo de elemento en caso de que haya conflictos, y seleccionar los elementos que se desean importar, rellenando los posibles campos que se deseen.
- Cumple con el requisito principal de utilidad, ahorrando gran cantidad de tiempo a aquellos investigadores que deban clonar en SICA la información que han introducido en otras plataformas de índole similar. Por lo general, un investigador debe introducir en diferentes plataformas dedicadas a la investigación la misma información, por lo que con esta herramienta se evita tener que volver a introducir los mismos datos en SICA, con el único requisito de poseer un fichero CVN con la información actualizada.

En definitiva, la herramienta desarrollada cubre los objetivos propuestos, proporcionando un sistema con la capacidad de importar un fichero CVN a la plataforma SICA, intentando obtener la mayor cantidad de información posible incluida en el currículum.

8.2. Líneas futuras

Aunque se dan por conseguidos los objetivos de este Trabajo de Fin de Máster, existen algunos aspectos que podrían mejorarse en un futuro, requerir un mantenimiento o implementarse adicionalmente.

Estas mejoras serían las siguientes:

- Mantener actualizada toda la información de campos y análisis de tráfico que han sido recopilados por SICA, ya que la página web podría cambiar su código HTML, añadir nuevos campos o alterar la forma en la que procesa los datos el servidor de la plataforma. En el caso de añadir nuevos campos que introduzca SICA en la plataforma la actualización de la herramienta no sería complicada siguiendo los pasos dados en este proyecto. Sin embargo, el cambio del código HTML (cambios de nombres de campos, cambio en la página a la que se envían los formularios, etc) o el cambio en el que procede el servidor actualmente serían bastantes problemáticos para la herramienta, ya que es

probable que algunos elementos dejen de poder ser importables. Por lo tanto, en estos últimos casos se requeriría una investigación para comprobar que ha cambiado exactamente y como se podría solventar el problema.

- Desarrollar la herramienta en otros lenguajes de programación, pudiendo hacer por ejemplo una aplicación móvil. Conociendo cómo funciona el formato CVN, el análisis de todos los elementos, campos y tráfico de SICA, no sería complicado replicar la herramienta que se ha desarrollado en otros lenguajes de programación, basándose en la herramienta web que se ha desarrollado.
- Permitir realizar otras operaciones desde la herramienta y que afecten a SICA, como por ejemplo, el borrado de elementos ya añadidos o la modificación de los mismos. En ambos casos sería necesario mostrar un listado de todos los elementos que han sido añadidos a SICA por el usuario, y dar la opción de borrar o modificar los mismos. Esto requiere realizar un nuevo análisis de tráfico con el objetivo de conocer el procedimiento para realizar estas acciones.
- Mejorar las funcionalidades a la hora de mostrar los campos de un determinado elemento en SICA. Existen diversos campos que no pueden ser mostrados al usuario dada su complejidad o irrelevancia, por lo que sería interesante estudiar estos campos y encontrar una manera para mostrarlos por pantalla.

Bibliografía

- [1] Ministerio de Economía, Industria y Competitividad. Currículum Vitae Normalizado (CVN). <https://cvn.fecyt.es/presentacion/que-es-cvn>, 2018.
- [2] Ministerio de Economía, Industria y Competitividad. Fundación Española para la Ciencia y la Tecnología. <https://www.fecyt.es/>, 2018.
- [3] Junta de Andalucía. Sistema de Información Científica de Andalucía (SICA). <https://sica2.cica.es>, 2018.
- [4] Junta de Andalucía. Centro Informático Científico de Andalucía (CICA). <https://www.cica.es/servicios/sistemas/sica2/>, 2018.
- [5] Biblioteca Nacional de Medicina de los Estados Unidos. PubMed. <https://www.ncbi.nlm.nih.gov/pubmed/>, 2018.
- [6] Thomson Reuters. ISI Web of Knowledge (WOK). <http://www.webofknowledge.com/>, 2018.
- [7] Elsevier. Scopus. <http://www.info.scopus.com/>, 2018.
- [8] Europass. Europass. <https://europass.cedefop.europa.eu>, 2018.
- [9] Jerome P McDonough. METS: standardized encoding for digital library objects. *International journal on digital libraries*, 6(2):148–158, 2006.
- [10] World Wide Web Consortium et al. SOAP Version 1.2. <http://www.w3.org/TR/soap12>, 2003.
- [11] Henriette D Avram. Machine-readable cataloging (MARC) program. *Encyclopedia of library and information science*, 3:1712, 2003.
- [12] Timothy D Jewell, Ivy Anderson, Adam Chandler, Sharon E Farb, Kimberly Parker, Nathan DM Robertson, et al. Electronic resource management. the report of the dlf initiative. 2004.

- [13] Tecnología e Innovación Productiva de Argentina Ministerio de Ciencia. Sistema Integral de Gestión y Evaluación. <http://www.mincyt.gob.ar/accion/cvar-6467>, 2018.
- [14] Consejo Nacional de Investigaciones Científicas y Técnicas. Sistema Integral de Gestión y Evaluación. <http://sigeva.conicet.gov.ar/>, 2018.
- [15] ORCID. ORCID. <https://orcid.org/>, 2018.
- [16] ResearcherID. ResearcherID. <http://www.researcherid.com/>, 2018.
- [17] Fundación Española para la Ciencia Y la Tecnología (FECYT). Manual de certificaciones de CVN. <https://cvn.fecyt.es/servicios-cvn/certificacion-cvn/procedimientos-de-validacion>, 2018.
- [18] Shudi Gao, C Michael Sperberg-McQueen, Henry S Thompson, Noah Mendelsohn, David Beech, and Murray Maloney. W3C XML schema definition language (XSD) 1.1 part 1: Structures. *W3C candidate recommendation*, 30(7.2):16, 2009.
- [19] PHP. ¿Qué es PHP? <http://php.net/manual/es/intro-what-is.php>, 2018.
- [20] Dalibor D Dvorski. Installing, configuring, and developing with Xampp. *Skills Canada*, 2007.
- [21] Juan Diego Gauchat. *El gran libro de HTML5, CSS3 y Javascript*. Marcombo, 2012.
- [22] Tim Bray, Jean Paoli, C Michael Sperberg-McQueen, Eve Maler, and François Yergeau. Extensible Markup Language (XML). *World Wide Web Journal*, 2(4):27–66, 1997.
- [23] Maximilian Hils. Mitmproxy. <https://mitmproxy.readthedocs.io/en/v2.0.2/introduction.html>, 2018.
- [24] S. Josefsson. The Base16, Base32, and Base64 Data Encodings. RFC 4648, RFC Editor, October 2006.
- [25] Ken Thompson. Programming techniques: Regular expression search algorithm. *Communications of the ACM*, 11(6):419–422, 1968.
- [26] James Garrett. Asynchronous JavaScript And XML. <http://www.adaptivepath.com/ideas/ajax-new-approach-web-applications/>, 2005.

-
- [27] David Cochran. *Twitter bootstrap web development how-to*. Packt Publishing Ltd, 2012.
 - [28] Douglas Crockford. The application/json media type for javascript object notation (json). Technical report, 2006.
 - [29] Ernesto Nebel and Larry Masinter. Form-based file upload in HTML. Technical report, 1995.
 - [30] Luke Welling and Laura Thomson. *PHP and MySQL Web development*. Sams Publishing, 2003.
 - [31] PHP. Sesiones en PHP.
<http://www.php.net/manual/es/session.examples.basic.php>, 2018.