

# UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS

Facultad 4



**Título: Plataforma Web para la Vigilancia Tecnológica**

**Autor:** Luyser Hernández Fariñas

**Tutor:** Dr. C. Omar Correa Madrigal

La Habana, diciembre 2022

“Año 64 de la Revolución”

# Dedicatoria

*A mi familia que tanto me ha apoyado y guiado a lo largo de todos estos años*

# Agradecimientos

*Primero que todo, quisiera agradecerles a mis padres, especialmente a mi padre, que, aunque no está presente para ver culminar esta importante etapa de mi vida, es gracias a él que he podido llegar hasta aquí.*

*A mis dos hermanas Anny y Katy con quienes siempre he podido contar en cualquier momento.*

*A mis amigos de toda la vida Carlos, Kyo, Isma, Mauro, Osmani y Fabián quien siempre han estado conmigo en las buenas y en las malas. A Lester quien se volvió mi mejor amigo dentro de la escuela y a Yill, que a pesar de haber llegado hace poco tiempo a mi vida, desde que lo hizo, me ha apoyado con el desarrollo de este trabajo.*

*No puedo dejar de mencionar a mi tutor Omar, a mi querida jefa de año Yadira, que puedo decir con toda la seguridad del mundo que ha sido la mejor profesora que he tenido a lo largo de mi vida como estudiante. Y a mi profesor y amigo Emilio que si no fuera por su ayuda y su apoyo no hubiera podido concluir esta tesis.*

*Por último, quisiera agradecerles a todas esas personas que han estado ahí para mí siempre que las he necesitado. Y aunque quisiera mencionarlas si lo hago no terminaría jamás.*

*En fin, muchas gracias a todos los que hicieron esto posible.*

# Declaración de Autoría

Declaramos ser autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales sobre esta, con carácter exclusivo.

Para que así conste firmamos la presente a los 1 días del mes de diciembre del año 2022.



Luyser Hernández Fariñas  
Autor



Dr.C.Omar Correa Madrigal  
Tutor



# Resumen

Con el fin de favorecer el desarrollo de la soberanía tecnológica, el Ministerio de las Comunicaciones en colaboración con la Universidad de las Ciencias Informáticas(UCI), ha decidido poner en marcha un proyecto para el desarrollo de una plataforma web para la vigilancia tecnológica, y su implementación en los observatorios tecnológicos. Estas plataformas realizan búsquedas especializadas a través de internet usando como base múltiples fuentes de información, para posteriormente realizar análisis y estudios sobre los resultados obtenidos. En este sentido, se presentan las bases de una herramienta que permita a los especialistas a realizar búsquedas profundas y eficientes. Para ello, se profundiza en los modos de interacción a través de interfaces visuales orientadas a la búsqueda avanzada y etiquetado de información por palabras claves.

**Palabras Clave:** Vigilancia tecnológica, recuperación de información, visualización de información, observatorio tecnológico, tecnología.

# Abstract

In order to promote the development of technological sovereignty, the Ministry of Communications in collaboration with the University of Informatics Sciences (UCI), has decided to launch a project for the development of a web platform for technological surveillance, and its implementation in technological observatories. These platforms carry out specialized searches through the Internet using multiple sources of information as a base, to later carry out analyzes and studies on the results obtained. In this sense, the bases of a tool that allows specialists to carry out deep and efficient searches are presented. To do this, the modes of interaction are deepened through visual interfaces oriented to advanced search and information labeling by keywords.

**Keyword's:** Technological surveillance, recover of information, information display, technological observatory, technology.

# ÍNDICE

<b>Introducción.....</b>	<b>10</b>
Problemática y Objetivos .....	15
<b>Capítulo 1 .....</b>	<b>17</b>
Introducción al Capítulo .....	17
1.1. Fundamentos de la Vigilancia Tecnológica .....	17
1.2. Fundamentos del Observatorio Tecnológico.....	18
1.3. Métodos y Técnicas Computacionales para la Búsqueda y Recolección de Información Científica .....	18
1.3.1. Fundamentos de la Recuperación de Información .....	18
1.3.2. Vocabulario controlado vs Vocabulario Controlado .....	22
1.3.3. Tecnología de Recuperación de Información .....	23
1.4. Estado Actual de los Sistemas de Vigilancia Tecnológica en Cuba ....	29
1.5. Tecnologías y Herramientas Utilizadas para dar Solución a la Problemática Propuesta .....	31
1.5.1. Metodologías de Software .....	32
1.5.2. Selección de Lenguaje .....	35
1.5.3. Selección de IDE .....	36
1.5.4. Django .....	38
Conclusiones Parciales.....	39
<b>Capítulo 2 .....</b>	<b>40</b>
Introducción al Capítulo .....	40
2.1 Descripción de la Propuesta de Solución.....	40
2.1.1 Requisitos Funcionales.....	40
2.1.2 Requisitos no Funcionales.....	41
2.2 Descripción de las Historia de Usuario.....	42
2.3 Plan de Iteraciones.....	46
2.4 Plan de Entrega .....	46
2.5 Estilo Arquitectónico.....	46
2.5.1 Arquitectura Cliente-Servidor.....	47
2.5.2 Patrón Arquitectónico .....	48
2.6 Tarjetas de Contenido, Responsabilidad y Colaboración .....	49
Conclusiones Parciales.....	50
<b>Capítulo 3 .....</b>	<b>50</b>

Introducción al Capítulo .....	50
3.1 Tareas de Ingeniería .....	50
3.1.1 Iteración I .....	51
3.1.2 Iteración II .....	52
3.1.3 Iteración III .....	54
3.1.4 Iteración IV .....	55
3.2 Descripción y Funcionamiento de la Plataforma .....	56
3.2.1 Página Principal.....	56
3.2.2 Página Resultado .....	57
3.3 Implementación.....	58
3.4 Pruebas de Aceptación .....	59
3.4.1 Iteración I.....	59
3.4.2 Iteración II.....	62
3.4.3 Iteración III.....	63
3.4.4 Iteración IV .....	64
Conclusiones Parciales.....	66
<b>Conclusiones.....</b>	<b>67</b>
<b>Bibliografía.....</b>	<b>68</b>

# Introducción

En la actualidad el mundo se ha visto sumergido completamente por la tecnología y el desarrollo continuo de la informática. Se puede comprobar fácilmente mirando, en cualquier dirección, hacia cualquier persona, institución, o área de la vida en sentido general, desde los teléfonos móviles que portan la gran mayoría de las personas hoy en día, hasta dispositivos en el hogar capaces de facilitar las labores de la misma. La tecnología y la informática se encuentran en todos lados, y constantemente se desarrollan nuevos dispositivos y algoritmos, que posibilitan el avance constante de la humanidad hacia un mundo cada vez más tecnológico.

Para las empresas, instituciones y organizaciones, que centran su labor en el desarrollo de la informática, la comunicación y las tecnologías para la sociedad, es necesario estar al corriente de cómo se comportan las mismas. Este proceso investigativo tiene que ser constante pues diariamente se realiza algún tipo de avance o innovación en este ámbito.

Para facilitar estos procesos en la actualidad existen sistemas para la Vigilancia Tecnológica(SVT). Sistemas por los que el país ha estado interesado desde hace muchos años y aunque se han intentado poner en ejecución algunos, se han visto limitados por distintos factores, lo que ha hecho que se terminen por usar plataformas internacionales lo cual limita la ejecución adecuada de la misma. A esto se suma que la vigilancia tecnológica no ha sido desarrollada ni generalizada de igual manera en todo el país.

Con el objetivo de lograr lo anteriormente planteado, el Ministerio de las Comunicaciones en colaboración con la UCI determina la necesidad de poner en práctica un Sistema de Vigilancia Tecnológica, el cual sea capaz de abarcar las distintas etapas del mismo, logrando una vigilancia completa y de calidad.

## Vigilancia Tecnológica

Con el fin de tener un mejor entendimiento de lo expuesto hasta el momento es necesario definir a que se refiere con el término Vigilancia Tecnológica. Se encuentra definida en la familia de normas UNE 166000, en concreto en la UNE 166000: *“Proceso organizado, selectivo y sistemático, para captar información del exterior y de la propia organización sobre ciencia y tecnología, seleccionarla, analizarla, difundirla y comunicarla, para convertirla en conocimiento con el fin de tomar decisiones con menor riesgo y poder anticiparse a los cambios”*

[1].

Según Steven C. Wheelwright, ex profesor emérito de la Universidad de Harvard, está constituida por el conjunto de técnicas que buscan organizar de forma sistemática la recolección, análisis y difusión de la explotación de informaciones técnicas útiles para el mantenimiento y el crecimiento de las empresas. *“consiste en prevenir y alertar a todo responsable de cualquier cambio, novedad o innovación ya sea técnica o científica. Teniendo en cuenta que puede cambiar el panorama o hacer ganar o perder una ventaja económica, la vigilancia se vuelve crítica y debe intervenir lo antes posible”*

[2].

Mientras que Andrés Manjarrés y Oswaldo Vanegas(2020) determinan que la VT es una herramienta para la integración y articulación de partes interesadas en un tema común, para encontrar posibles soluciones en los avances y las tendencias tecnológicas dadas en el país y el resto del mundo.

Debido a la gran importancia de la implementación de un Sistema de Vigilancia Tecnológica, se propone para la utilización en los Observatorios Tecnológicos, la implementación de una aplicación web que logre desplazar a las plataformas extranjeras para VT que se utilizan en nuestro país. Creando de este modo una plataforma que posibilite a los especialistas del observatorio tecnológico llevar a cabo un adecuado análisis competitivo de las tendencias y avances en las tecnologías de la información y las comunicaciones

## Etapas clave del ciclo de vigilancia tecnológica e inteligencia

Existe un consenso generalizado sobre las etapas que ha de seguir el proceso de vigilancia e inteligencia tecnológica, denominado ciclo de vigilancia e inteligencia. En esencia, este ciclo pretende dar respuesta a unas cuestiones fundamentales:



Imagen 1. Ciclo de la vigilancia tecnológica [3].

1. **Identificar** y analizar las necesidades de información.

Se identifican las necesidades de información, las tecnologías a vigilar y los factores críticos de vigilancia (FCV) para definir la estrategia de vigilancia más adecuada para la organización. Los factores críticos de vigilancia son los aspectos claves a vigilar y se determinan por cada actividad de la cadena de valor de la organización. Además, para precisar la búsqueda de información, es recomendable acompañar los FCV con descriptores, palabras clave, prioridad, horizonte temporal, etc.

2. **Buscar** información.

Para ello, se definen los objetivos de la búsqueda de información y se elabora la estrategia para precisar las necesidades, localizar la información y capturarla de una manera organizada. Para lo cual se realizan tareas como: identificación de

palabras clave, validación de expertos, selección de fuentes de información relevantes y formulación de ecuaciones de búsqueda.

3. **Analizar**, procesar y valorizar la información.

Para ello, se combinan criterios de validación de la información obtenida, técnicas analíticas de información y herramientas informáticas especializadas que ayuden al equipo a seleccionar la información relevante según los objetivos de búsqueda. En esta fase, resultan de ayuda herramientas como: mapas tecnológicos, software de patentes, gestores bibliográficos, visualización de información o software integrales de vigilancia tecnológica.

4. **Valorizar** los resultados.

Para ello, una vez concretados los resultados y valorada su trascendencia, se han de generar los denominados productos de vigilancia. Estos son soportes de información confeccionados con los resultados de información obtenidos del proceso de vigilancia, y que conformarán el medio de difusión de éstos en la organización.

Los más relevantes son:

- Boletín o informe de vigilancia tecnológica.
- Boletín de oportunidades tecnológicas.
- Informes de prospectiva y tendencias.
- Estudios de patentes.
- Seguimiento del entorno.

5. **Difundir** resultados

Diseminar los resultados del proceso a las personas con responsabilidades pertinentes en la organización. Para ello, se ha de diseñar una estrategia de comunicación interna eficaz y distribuida en toda la organización, que cubra las necesidades de información del personal y utilice los medios de comunicación más generalizados en la organización, abarcando tanto los informales como los formales.

6. **Orientar** toma de decisiones

Apoyar el proceso de toma de decisiones de los empleados en la organización. Para ello, a partir de la difusión de los productos de vigilancia tecnológica, se debe promover la reflexión interna y colectiva sobre a las implicaciones tecnológicas, productivas y competitivas de los resultados obtenidos.

Este al ser un proceso tan largo, se ha decidido separarlo en distintas etapas. Este trabajo estará orientado a la primera etapa del mismo el cual comienza una vez se realizada la identificación de la información a la cual se le desea hacer seguimiento, por medio de los expertos. Una vez identificados se inicia el proceso en la segunda cuestión: Búsqueda y recopilación de la información.

## **Plataformas Cubanas para la Vigilancia Tecnológica**

En Cuba, desde hace muchos años se han intentado llevar a cabo el desarrollo de herramientas de recuperación de información, con especial énfasis en la información científica. Un claro ejemplo es Infomed, que constituye toda una red al servicio de la Salud que provee diferentes herramientas para la búsqueda e intercambio de información científica. Existen también otras organizaciones como: Biomundi (perteneciente al IDICT, Ministerio de Ciencia Tecnología y Medio Ambiente), Oficina Cubana de Propiedad Intelectual (OCPI, CITMA), Centro Nacional de Investigaciones Científicas (CNIC, perteneciente al Ministerio de Educación Superior), DELFOS (Ministerio de Informática y Comunicaciones) que desde hace más de 15 años practican la VT, no solo para sí, sino como parte de su objeto social, a terceros. [4].

Existen además muchas bibliotecas y repositorios cubanos que ayudan a satisfacer la necesidad de consumo de información. Se han realizado investigaciones y lanzado artículos científicos que proponen tecnologías para la búsqueda de documentación basada en una visualización de árbol, y que brinda los datos bibliográficos de los artículos buscados, para posteriormente poder realizar una búsqueda del mismo, para un análisis detallado. A pesar de las diversas investigaciones y artículos que abordan la necesidad de crear sistemas de VT, se han logrado implementar muy pocas y estas que se han puesto en marcha son muy básicas y no logran alcanzar todas las etapas de la VT. Esto ha obligado a las empresas a migrar hacia plataformas extranjeras que les posibilitan llevar a cabo una adecuada vigilancia. Aunque en sus inicios muchas de estas plataformas ofrecían sus servicios de modo gratuitos, en la actualidad para poder usarlas es necesario pagar licencias, lo cual constituye una fuerte limitante. Algunas de estas plataformas utilizadas son: Xerka, Vigiale, Vixia, Vicubo, SoftVT, Miniera, Cosmos y Hontza, esta última siendo una de las mejores y más completas plataformas de VT utilizada principalmente en Cuba por la empresa Biocubafarma. De esta manera los procesos búsqueda y recopilación de información científica se ven seriamente afectados y tardan mucho tiempo [17]

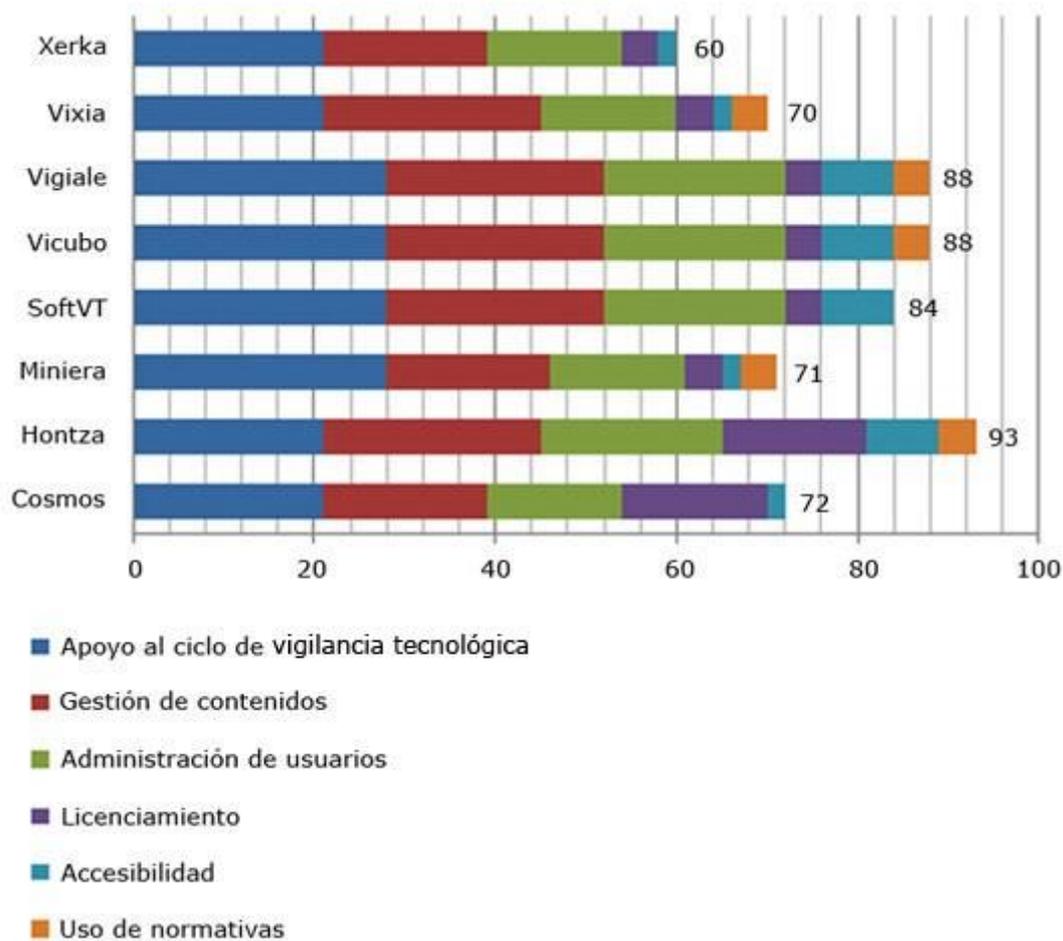


Imagen 1. Comparación de distintas plataformas de VT teniendo en cuenta distintos indicadores [5] [17]

## Métodos de investigación

Para llevar a cabo la investigación se determinó el uso de dos métodos de investigación. Los métodos seleccionados pertenecen al marco empírico de la metodología de la investigación: “la entrevista” y “las encuestas”.

Las entrevistas constituirán la base fundamental para la realización; la realización de las mismas posibilita la interacción directa con los involucrados en el proceso que se va a desarrollar, permitiendo de este modo obtener la información de manera directa y precisa, posibilitando un enfoque más coherente de las acciones a llevar a cabo para el desarrollo de la solución. Para poder llevar a cabo este método se proponen una serie de preguntas, que permiten conocer cómo se comporta la situación problemática en la institución, y de este modo poder determinar una solución factible.

- ¿De qué manera se efectúan los procesos de Vigilancia Tecnológica en la institución?
- De existir procesos de Vigilancia Tecnológica ¿Que trabas existen en la ejecución de los mismos?
- ¿Bajo qué normas se sustentan para determinar la factibilidad de la información obtenida?

- ¿Cuáles son las principales herramientas utilizadas por la institución para la recopilación de información, para un posterior análisis?
- De utilizar alguna herramienta muy específica ¿Su uso genera algún tipo de problema? ¿Cuáles son los principales problemas que genera?
- ¿La utilización de dicha herramienta genera alguna especie de costo?
- ¿Qué cambios deberían tener estas herramientas para mejorar su funcionamiento?

En el caso de las encuestas, juegan un papel crucial a la hora de evaluar el sistema que se desea desarrollar. Mediante esta se selecciona un conjunto de usuarios se enfrentarán a la aplicación para que la pongan a prueba, enfrentando la aplicación a un entorno real, posteriormente se les brindara una encuesta, la cual permitirá a los desarrolladores conocer la opinión de los usuarios con respecto al sistema, y así conocer los principales problemas que presenta la plataforma, así como posibles mejoras en cuanto a diseño y funcionalidad que permitan mejorar la experiencia del usuario. De este modo se garantiza una solución óptima y es una garantía del cumplimiento de todos los objetivos planteados.

## **Problemática y Objetivos**

Tomando en cuenta todo lo anteriormente planteado se llegar a determinar que el problema que se presenta es ¿Cómo contribuir a la búsqueda y recopilación de información científica para favorecer a los procesos de vigilancia tecnológica en el país?

**El objeto de estudio** está centrado en los métodos y técnicas computacionales para la búsqueda y recolección de información científica.

### **Campo de acción:**

El campo de acción se encontraría enmarcado en los métodos y técnicas computacionales para la Búsqueda y Recolección de publicaciones científicas.

### **Objetivo general:**

Desarrollar un sistema que contribuya a la búsqueda y recopilación de información científica para favorecer a los procesos de vigilancia tecnológica en el país.

### **Tareas de Investigación:**

Partiendo del objetivo general anteriormente planteado se llevarán a cabo las siguientes tareas de investigación:

1. Realizar un análisis del estado del arte de las distintos métodos y herramientas de software especializadas en la vigilancia tecnológica, prestando mayor interés a todas aquellas que se especialicen en las etapas de búsqueda y recolección de información.

2. Desarrollo de un sistema de búsqueda y recolección de publicaciones científica utilizando como principal medio de obtención de la misma un vocabulario controlado.
3. Realizar las pruebas de software pertinentes el sistema.

# Capítulo 1

“Fundamentación teórica”

## Introducción al Capítulo

A lo largo de este capítulo se expondrán los conceptos necesarios para facilitar la comprensión del tema desarrollado. Además, se determinan y se describirán las distintas herramientas y metodologías que permitirán llegar a lograr la propuesta de solución. El presente capítulo cuenta con 5 epígrafes los cuales se encuentran divididos en sub-epígrafes. Los distintos temas recogidos en cada uno son: Fundamentos de la Vigilancia Tecnológica (epígrafe 1.1), Fundamentos del Observatorio Tecnológico (epígrafe 1.2), Métodos y Técnicas Computacionales para la Métodos y Técnicas Computacionales para la Búsqueda y Recolección de Información Científica (epígrafe 1.3) Estado Actual de los Sistemas de Vigilancia Tecnológica en Cuba (epígrafe 1.4), Tecnologías y Herramientas Utilizadas para dar Solución a la Problemática Propuesta (1.5)

### 1.1. Fundamentos de la Vigilancia Tecnológica

El desarrollo de un sistema para la búsqueda y recolección de publicaciones científicas, tiene como objetivo principal lograr un mayor avance en los procesos de Vigilancia Tecnológica que se efectúen en cualquier institución del país. Por este motivo es necesario plantear los conceptos y fundamentos que lleven a alcanzar un mejor entendimiento de los sistemas de vigilancia tecnológica.

Para María I. Ramírez, David Escobar y Bibiana Arango, el objetivo principal de la Vigilancia Tecnológica es el proceso diario de conocer todo lo que ocurre en un área tecnológica determinada, con el fin de trazar el camino que deben seguir los trabajos de investigación desarrollados dentro de una organización y de esta manera lograr discernir, pero con conocimiento.

Según la norma UNE 166006:2011 Ex Gestión de la I+D+i: Sistema de Vigilancia Tecnológica e Inteligencia Competitiva, la Vigilancia Tecnológica es un proceso organizado, selectivo y permanente, de captar información del exterior y de la propia organización sobre ciencia y tecnología, seleccionarla, analizarla, difundirla y comunicarla, para convertirla en conocimiento para tomar decisiones con menor riesgo y poder anticiparse a los cambios.

Se puede resumir entonces que la Vigilancia Tecnológica es el proceso de recapitulación de la información, para posteriormente realizarle un análisis y tomar decisiones con el objetivo anticiparse a los cambios que puedan ocurrir y en el término empresarial poder mantenerse delante de la competencia.

Ya anteriormente se hablaba de la gran importancia que representa la Vigilancia Tecnológica para todas aquellas instituciones ya sean empresas o universidades. Pero

a la hora de emprender esta búsqueda de información ¿cuáles son las fuentes que se deben tener en cuenta? Para realizar un estudio de este tipo lo ideal es utilizar publicaciones científicas, pues las mismas constituyen las principales fuentes de información en el campo de las ciencias, y es de este modo que se divulgan investigaciones realizadas y hallazgos obtenidos [18] [19]

## **1.2. Fundamentos del Observatorio Tecnológico**

Pero no se puede hablar de los procesos de vigilancia tecnológica sin hablar de los observatorios tecnológicos. El mismo, es un espacio que permite gestionar en tiempo real cuales son las últimas noticias, últimas publicaciones o investigaciones científicas realizadas, en este caso en la esfera de la informática y dentro de esta en un área determinada. Permite estar al corriente del comportamiento a nivel mundial de la tecnología y por lo tanto permite estar en la vanguardia. Se puede apreciar no solo desde el punto de vista del software, sino también desde el hardware. Todo condicionado alrededor de la información que se desea buscar y analizar. [20] [21]

“Los observatorios tecnológicos, para mantener una adecuada vigilancia tecnológica, implementan métodos de recuperación de la información entre los que se encuentra el uso de las API y el web scraping”. Por medio de esta frase, y de los conceptos anteriormente planteados, se puede llegar a tener una visión general del proyecto a desarrollar. Permitiendo comprender el funcionamiento de las herramientas fundamentales y las ventajas que las mismas pueden ofrecer.

## **1.3. Métodos y Técnicas Computacionales para la Búsqueda y Recolección de Información Científica**

Son todos aquellos procesos destinados a la recuperación de información, desde la generación de las colecciones, su depuración, indexado, tratamiento textual, clasificación, almacenamiento, recuperación mediante modelos booleanos, vectoriales, probabilísticos, basados en el lenguaje, así como todos aquellos elementos que inciden en cualquier aspecto relacionado como por ejemplo el interfaz de consulta, el comportamiento del usuario, la retroalimentación de las consultas y la representación de la información.

La recuperación de información no puede concebirse como un elemento único e insoluble como por ejemplo única y exclusivamente un modelo o algoritmo de recuperación. La recuperación de información debe estudiarse como un compendio de procesos altamente interrelacionados que conforman una verdadera cadena documental en recuperación de información.

[22]

### **1.3.1. Fundamentos de la Recuperación de Información**

La Recuperación de Información es parte de la informática que estudia la recuperación de la información (no datos) de una colección de documentos escritos. Los documentos recuperados pueden satisfacer una necesidad de información de un usuario expresada normalmente en lenguaje natural [22]

La recuperación de información se centra en la representación, almacenamiento, organización y acceso a elementos de información. Estos procesos deberían proporcionar al usuario la capacidad de acceder a la información que necesita. Sin embargo, existe un problema bastante importante en lo referente a la caracterización de las necesidades de información del usuario, que no suele ser fácil solucionar. Procesa archivos de registros y peticiones de información, e identifica y recupera de los archivos ciertos registros en respuesta a las peticiones de información. [23] [25]

Es el proceso por el cual las demandas informativas y documentales del usuario son resueltas en un sistema de información, compuesto por un corpus documental de volumen variable, cuyo tratamiento de indexación y almacenamiento hacen posible su estructuración, interrogación y representación, por medio del empleo de algoritmos matemáticos, estadísticos y semánticos. [22]

### **La consulta:**

**-Necesidad de información:** Es la declaración en lenguaje natural de la información que demanda o requiere el usuario para el desempeño de sus actividades y funciones.

**-Formulación del usuario:** Proceso racional del usuario para confeccionar la frase o sucesión de términos con los que efectuará la consulta o interrogación del sistema.

**-Consulta del usuario:** Es la expresión con la que se configura la demanda informativa del usuario, por regla general, en lenguaje natural, utilizando los términos y palabras que le resultan más aproximados al objeto de recuperación o cuya previsión e intuición le sugieren un mejor aprovechamiento del sistema en su búsqueda

**-Formulación del sistema:** Procesamiento y reformulación de la consulta del usuario que implica su descomposición en unidades mínimas (término a término), sustitución de caracteres extraños, procesos de reducción, eliminación de palabras vacías, eliminación de signos diacríticos, sustitución y adición de términos normalizados. Finalmente, una vez depurada y adaptada la consulta, se aplican los operadores booleanos y especiales propios del algoritmo de recuperación que se fuere a emplear

**-Consulta del sistema:** Es el resultado de la formulación del sistema partiendo de la consulta del usuario. Por regla general una sentencia de consulta optimizada para la recuperación en el sistema de información que equivale a la expresada por el usuario en lenguaje natural. Dicho de otra forma, es la traducción de la consulta del usuario a un lenguaje documental o técnico, propio de la recuperación de información.

**-Expansión de consulta:** Es un proceso de reformulación automática del sistema que permite añadir nuevos términos a la consulta para mejorar el contexto de la consulta original del usuario. Esto se consigue mediante procesos de clustering, que determinan la frecuencia de aparición de un grupo de términos contiguos, relacionados con la consulta del usuario, presentes en documentos clasificados dentro de un mismo ámbito temático (en el caso de análisis del contexto local) y en torno a toda la colección (en el caso de análisis del contexto global)

**-Patrón:** Expresión sintáctica que define una serie de caracteres textuales, alfabéticos, numéricos y especiales, que se ajustarán por coincidencia en una palabra o término de un texto determinado.

**-Expresión regular:** También conocidas como REGEXP y POSIX, son aquellas expresiones sintácticas complejas y normalizadas, compuestas a base de patrones que permiten la definición de consultas de datos en un corpus documental dado, mediante cadenas de caracteres, repeticiones y concatenaciones, establecidas por sus reglas de construcción. [26].

### **La base del conocimiento:**

**-Colección:** Es sinónimo de base de conocimiento, fondo, biblioteca de documentos o corpus documental. El concepto colección hace referencia a un compendio de documentos seleccionados previamente u obtenidos mediante métodos de minería de datos.

**-Colección de referencia:** Aquella colección utilizada para la experimentación de los modelos de recuperación de información y sus algoritmos. Ello implica la disposición de plantillas de resultados con los documentos relevantes para cada consulta de prueba, de cara a la evaluación del SRI.

**-Documento:** Elemento básico con el que se conforman las colecciones y unidad básica de recuperación. Se considera documento a todo tipo de información independiente, artículos, monografías, sitios y páginas web, resúmenes, textos completos, etc.

**-Documento sustituto:** Símil de un documento de una colección, fiel a sus contenidos mediante sus elementos básicos como título, resumen, frase de contextualización y URL. Se utiliza en las páginas de resultados, en procesos de visualización y representación.

### **Evaluación y resultados de la recuperación:**

**-Precisión:** En recuperación de información, precisión es la medida que define cuantitativamente la relación entre los documentos recuperados y su relevancia para satisfacer la consulta del usuario.

**-Exhaustividad:** También denominado Recall es la capacidad del sistema de información para recuperar todos los documentos relevantes con respecto a la totalidad de los existentes en la colección, de acuerdo a los condicionamientos y especificaciones de la consulta del usuario.

**-Pertinencia:** Aquel documento que añade nueva información a la previamente almacenada en la mente del usuario, que le resulta útil en el trabajo que ha propiciado la pregunta. El conjunto pertinente de documentos recuperados puede definirse como el subconjunto de los documentos almacenados en el sistema que es apropiado para la necesidad de información del usuario. [27]

**-Relevancia:** Un mismo documento puede ser considerado relevante, o no relevante, por dos personas distintas en función de los motivos que producen la necesidad de información o del grado de conocimiento que sobre la materia posean

ambos. Llegados a un caso extremo, un mismo documento puede parecer relevante o no a la misma persona en momentos diferentes de tiempo. Aunque puede usarse otra terminología, la voz relevancia parece la más apropiada para indicar la relación entre un documento y una petición de información efectuada por un usuario, aunque puede resultar erróneo asumir que ese grado de relación es fijo e invariable, siendo mejor decir, que un documento ha sido juzgado como relevante a una específica petición de información. Es el grado de importancia y significación que concede el usuario a los resultados obtenidos en un sistema de información. [22], [27]

**-Rendimiento:** Es un factor para la evaluación de un sistema de recuperación de información, que se obtiene evaluando la pertinencia y exhaustividad de los resultados generados por un conjunto de consultas de prueba en la colección de referencia, con respecto a las soluciones propuestas para el mismo por los especialistas.

**-Ranking:** También denominado alineado de los documentos es el proceso de evaluación de los resultados obtenidos, tras aplicar un modelo de recuperación de información, reflejando en un coeficiente o indicador numérico la relevancia, precisión y exhaustividad de los mismos, para una consulta dada por el usuario.

## **Sistemas de recuperación de información**

**-Tarea:** de recuperación. Aquellas rutinas algorítmicas ejecutadas por el sistema de información en respuesta a una solicitud del usuario.

**-Algoritmo:** de recuperación. Es el conjunto de métodos documentales, rutinas de tratamiento de información y procedimientos automáticos de tipo matemático-estadístico, ya predefinido en el funcionamiento de un programa informático, tales como la depuración, indexación, comparación de consultas, aplicación de modelos de recuperación, representación, evaluación y análisis necesarios para que el sistema de información satisfaga las necesidades de información del usuario. El orden en que se ejecutan, la casuística de la consulta y la experiencia del usuario, son factores que influyen en la ejecución de los algoritmos de recuperación, generando un grado de variabilidad en los resultados obtenidos.

**-Filtrado:** Proceso de refinamiento y perfección de la consulta del usuario por el que se delimita, especifica o amplía la búsqueda original, una vez que los resultados de la búsqueda satisfacen parcialmente la demanda informativa del usuario.

**-Coincidencia exacta:** Es el mecanismo por el cual sólo los documentos que satisfacen algunos criterios y rasgos bien especificados en la consulta son recuperados y devueltos al usuario como una respuesta unívoca, cumpliéndose al 100% en sus expectativas.

**-Recuperación de datos:** La recuperación de elementos (tuplas, los objetos, páginas Web, documentos) cuyo contenido cumple los requisitos especificados en una consulta de usuario basada en expresión regular o por coincidencia de patrones [23].

Para que el sistema a desarrollar alcance el objetivo deseado, el mismo debe ser capaz de realizar la recuperación de la información de forma óptima mediante un adecuado control de las consultas. La utilización de distintas fuentes de información que sean capaces de satisfacer las necesidades que se pretenden sustentar. El sistema debe tener la capacidad de asignar una evaluación de la información obtenida, teniendo algunos parámetros como, por ejemplo: mientras mayor número de citas posea un

artículo obtenido, mayor será la evaluación. También darle la oportunidad al usuario de subirle la evaluación a algún artículo. De este modo garantizando que el producto alcance una alta eficiencia.

### **1.3.2. Vocabulario controlado vs Vocabulario Controlado**

Para garantizar la elaboración de un buen trabajo académico hay que garantizar como paso inicial la elaboración de una estrategia de búsqueda de información. Esto se debe a que de esta manera se asegura que la información obtenida para desarrollar la investigación es la mejor y más adecuada. Para lograrlo es fundamental la utilización de un vocabulario compuesto por los conceptos necesarios, para esto primero hay que determinar qué tipo de vocabulario es mejor, uno libre o uno controlado.

#### **Vocabulario controlado:**

Está compuesto por un índice de términos normalizados dentro de un lenguaje documental llamados descriptores que representan de manera unívoca un concepto, homogeneizando los términos empleados para una búsqueda más precisa. De este modo, agrupamos sinónimos con un solo término y se tiene un mayor control de las polisemias, variantes ortográficas etc... Por otro lado, el uso de un vocabulario controlado facilita la buena elección de los términos adecuados para nuestra estrategia de recuperación de información pues nos proporcionan: definiciones de los mismos, relaciones con otros términos (jerarquías, de equivalencia y asociativas) y en muchas ocasiones la traducción de dichos términos a varios idiomas.

#### **Vocabulario libre:**

Es el utilizado en la comunicación humana por lo que tiene como ventaja que es ilimitado dando a lugar a una mayor abundancia de vocablos, polisemias..., que a su vez producen ambigüedad e inexactitud en los resultados que podamos obtener en nuestra recuperación de información [28].

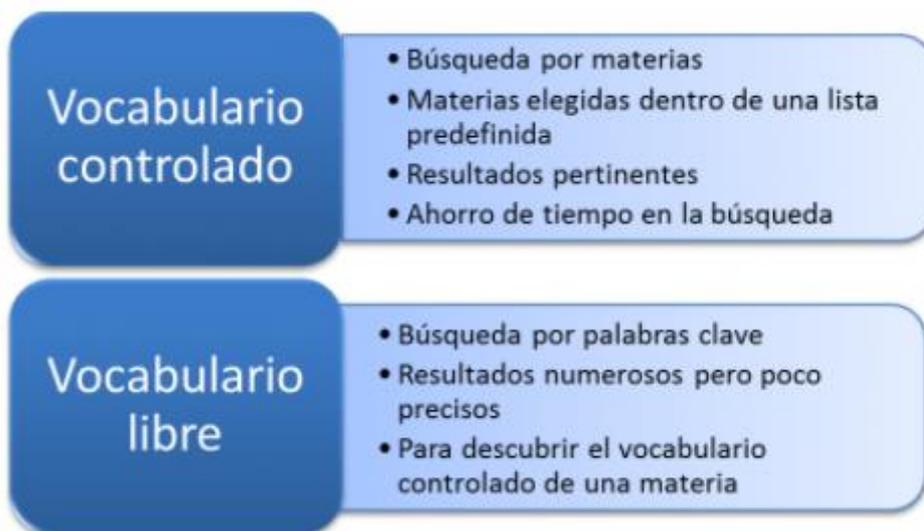


Imagen 2. Comparación entre Vocabulario Controlado y Vocabulario Libre [28].

Una vez comprendido lo anteriormente expuesto. Se determina que para la realización del sistema se utilizara un vocabulario controlado. Para la elaboración del mismo se utilizará como base “Computer Science Ontology”, en español “Ontología de la Ciencia de la Computación”, o por sus siglas en inglés “CSO”. Con el fin de apoyar a los usuarios con experiencia en el uso de esta aplicación, se utilizará una barra de búsqueda que solo permitirá introducir como parámetro de búsqueda palabras referente a los temas encontrados en el árbol.

### 1.3.3. Tecnología de Recuperación de Información

Para el desarrollo de la plataforma es necesario implementar distintos métodos de búsqueda de información, para poder llevar a cabo el propósito principal de la misma. Los principales métodos a utilizar son las API, los canales RSS, y en un último caso se utilizaría la integración de las API con el webscraping. Los dos primeros métodos de obtención permiten obtener mediante ciertos parámetros la información deseada de determinado sitio web.

Las API facilitan en gran medida el desarrollo de una aplicación, esto se debe a que permiten la reutilización de funcionalidades que se encuentren en otro sitio web, sin necesidad de que el desarrollador las implemente. En la plataforma a desarrollar gracias a la utilización de las API, se pueden realizar búsquedas en distintos sitios tan solo enviando la palabra clave y en determinados casos una apikey única, que es la que permite obtener la información.

El funcionamiento de los canales RSS es incluso más simples, estos con tan solo establecen la comunicación con ellos, y mostrando la información que dichos canales pueden compartir y se solicita

En el caso del web scraping es muy útil en el desarrollo de la herramienta pues esta ingresa a determinada url y permite obtener toda la información perteneciente a la misma, pero en este caso queda como último recurso, se le da más prioridad a la obtención de la información mediante, pues un gran número de páginas web tienen bots anti scraping pues este método se utiliza para robar información.

### **Api:**

Las *Application Programming Interfaces*, más conocidas como API, como su nombre lo indica, son interfaces de programación de aplicaciones. Son un puente de comunicación que permiten consumir los recursos ofrecidos por determinadas aplicaciones, en este caso páginas web y utilizarlos para el desarrollo otras aplicaciones. Ofrecen a los desarrolladores agilidad y rapidez al poder integrar en muy poco tiempo funcionalidades sin necesidad de crearlas desde cero, disminuyendo el tiempo de desarrollo. Las API ya han sido probadas en múltiples ambientes de desarrollos asegurando un óptimo desempeño en cualquiera de los ambientes en los que se utilice siendo así una garantía de calidad. Debido a la competencia que existente en la web las compañías están en constante creación de nuevas funcionalidades y mejoras de los servicios que ofrecen, mejoras que se le ofrecen a terceros mediante las API, estando así en un proceso de mejora continua [29] [30].

### **Desventajas de la utilización de Api:**

Así como las API ofrecen grandes ventajas, también traen consigo algunas desventajas. Las aplicaciones web muchas veces sufren actualizaciones periódicas. Estas se hacen con el objetivo de agregar nuevas funcionalidades, hacer mejoras a las ya existentes o eliminar algunas que por algún motivo ya no se utilizan, que poseen problemas, o simplemente por cuestiones de seguridad. Como mismo ocurre esto, las API de un software también sufren cambios y puede ocurrir que a la hora de establecer la comunicación se haya perdido el acceso a la API, o incluso puede ocurrir que en un principio el acceso a la misma haya sido gratuito, pero posteriormente los desarrolladores deciden que el acceso a este servicio deberá ser pago. Al consumir las funcionalidades de que ofrece la aplicación de un tercero, como se comentaba anteriormente ayuda en el desarrollo de un software propio, pero al no ser propio, el desarrollador no tiene control sobre dichas funcionalidades, y si estas dejan de funcionar de manera repentina, no se puede dar una solución al mismo, solamente quedaría esperar que los desarrolladores de la API le den solución. Al momento de querer implementar por primera vez un servicio de tercero como las API, la falta de manejo y de conocimiento puede afectar el desempeño pues no todas son fáciles de implementar y no todas tienen los mismos requerimientos para su funcionamiento [29] [30].

### **API REST:**

REST [30] [31] no es un protocolo ni un estándar, sino más bien un conjunto de límites de arquitectura. Los desarrolladores de las API pueden implementarlo de distintas maneras.

Cuando el cliente envía una solicitud a través de una API de RESTful, esta transfiere una representación del estado del recurso requerido a quien lo haya solicitado o al extremo. La información se entrega por medio de HTTP en uno de estos formatos: JSON (JavaScript Object Notation), HTML, XML, Python, PHP o texto sin formato. JSON es el lenguaje de programación más popular, ya que tanto las máquinas como las personas lo pueden comprender y no depende de ningún lenguaje, a pesar de que su nombre indique lo contrario.

También es necesario tener en cuenta otros aspectos. Los encabezados y los parámetros también son importantes en los métodos HTTP de una solicitud HTTP de la API de RESTful, ya que contienen información de identificación importante con respecto a los metadatos, la autorización, el identificador uniforme de recursos (URI), el almacenamiento en caché, las cookies y otros elementos de la solicitud. Hay encabezados de solicitud y de respuesta, pero cada uno tiene sus propios códigos de estado e información de conexión HTTP.

Para que una API se considere de RESTful, debe cumplir los siguientes criterios:

- Arquitectura cliente-servidor compuesta de clientes, servidores y recursos, con la gestión de solicitudes a través de HTTP.
- Comunicación entre el cliente y el servidor sin estado, lo cual implica que no se almacena la información del cliente entre las solicitudes de GET y que cada una de ellas es independiente y está desconectada del resto.
- Datos que pueden almacenarse en caché y optimizan las interacciones entre el cliente y el servidor.
- Una interfaz uniforme entre los elementos, para que la información se transfiera de forma estandarizada. Para ello deben cumplirse las siguientes condiciones:
  - Los recursos solicitados deben ser identificables e independientes de las representaciones enviadas al cliente.
  - El cliente debe poder manipular los recursos a través de la representación que recibe, ya que esta contiene suficiente información para permitirlo.
  - Los mensajes auto descriptivos que se envíen al cliente deben contener la información necesaria para describir cómo debe procesarla.
  - Debe contener hipertexto o hipermedias, lo cual significa que cuando el cliente acceda a algún recurso, debe poder utilizar hipervínculos para buscar las demás acciones que se encuentren disponibles en ese momento.
- Un sistema en capas que organiza en jerarquías invisibles para el cliente cada uno de los servidores (los encargados de la seguridad, del equilibrio de carga, etc.) que participan en la recuperación de la información solicitada.
- Código disponible según se solicite (opcional), es decir, la capacidad para enviar códigos ejecutables del servidor al cliente cuando se requiera, lo cual amplía las funciones del cliente.

Si bien la API de REST debe cumplir todos estos parámetros, resulta más fácil de usar que un protocolo definido previamente, como SOAP (protocolo simple de acceso a objetos), el cual tiene requisitos específicos, como la mensajería XML y la seguridad y el cumplimiento integrados de las operaciones, que lo hacen más lento y pesado.

Por el contrario, REST es un conjunto de pautas que pueden implementarse según sea necesario. Por esta razón, las API de REST son más rápidas y ligeras, cuentan con mayor capacidad de ajuste y, por ende, resultan ideales para el Internet de las cosas (IoT) y el desarrollo de aplicaciones para dispositivos móviles.

### **API SOAP:**

SOAP [31] es un protocolo estándar que se creó originalmente para permitir la comunicación entre las aplicaciones que se diseñaban con diferentes lenguajes y en diferentes plataformas. Como es un protocolo, impone reglas integradas que aumentan la complejidad y la sobrecarga, lo cual puede retrasar el tiempo que tardan las páginas en cargarse. Sin embargo, estos estándares también ofrecen normas integradas que pueden ser ideales para el sector empresarial. Los estándares de cumplimiento integrados incluyen la seguridad, la atomicidad, la uniformidad, el aislamiento y la durabilidad (ACID), que forman un conjunto de propiedades que garantizan operaciones confiables de las bases de datos.

Las especificaciones comunes de los servicios web incluyen lo siguiente:

- **Seguridad de los servicios web (WS-Security):** estandariza la forma de proteger y transferir los mensajes usando identificadores únicos llamados tokens.
- **Mensajería segura de los servicios web (WS-ReliableMessaging):** estandariza el control de errores entre mensajes que se transfieren en infraestructuras de TI poco confiables.
- **Abordaje de los servicios web (WS-Addressing):** paquetes que enrutan la información como metadatos dentro de los encabezados SOAP, en lugar de mantener la información en un lugar más profundo de la red.
- **Lenguaje de descripción de los servicios web (WSDL):** describe qué hace un servicio web, así como dónde comienza y termina.

El envío de una solicitud de datos a una API de SOAP se puede administrar a través de cualquiera de los protocolos de la capa de la aplicación: HTTP (para los exploradores web), SMTP (para el correo electrónico), TCP, entre otros. Sin embargo, una vez que se recibe una solicitud, los mensajes SOAP de retorno deben ser documentos XML, que es un lenguaje de marcado que comprenden las personas y las máquinas. Una solicitud completa a una API de SOAP no se almacena en caché por un navegador, por lo que no se puede acceder a ella después sin reenviarla a la API.

### **Diferencias entre SOAP y REST:**

Es posible que muchos sistemas heredados sigan rigiéndose por SOAP, aunque REST haya surgido más tarde y se considere una alternativa más rápida en los escenarios basados en la Web. REST es un conjunto de pautas que ofrece una implementación flexible, mientras que SOAP es un protocolo con requisitos específicos, como en el caso de la mensajería XML [32].

Las API de REST son ligeras, así que son ideales para los contextos más nuevos, como el Internet de las cosas (IoT), el desarrollo de aplicaciones móviles y la informática sin servidor. Los servicios web de SOAP ofrecen seguridad y cumplimiento de las operaciones integradas que coinciden con muchas de las necesidades empresariales, pero que también los hacen más pesados. Asimismo, muchas API públicas, como la API de Google Maps, siguen las pautas de REST [30].

## **Canales RSS o Feeds RSS:**

El sistema RSS está actualmente en la versión 2.0, y el acrónimo es una abreviatura del término Really Simple Syndication (algo como "distribución muy simple"). Sus versiones anteriores se llamaban Rich Site Summary (RSS 0.91) y RDF Site Summary (RSS 0.9 y 1.0).

Feeds RSS funcionan con un archivo XML conocido como webfeed, Atom o, simplemente, RSS feed. Es dentro de este archivo donde se encuentran los datos relacionados con el contenido del sitio, como la descripción, el autor, las imágenes y los enlaces de destino.

Un **feed RSS** es un archivo generado por algunos sitios web (y por muchos weblogs) que contiene una versión específica de la información publicada en esa web. Cada elemento de información contenido dentro de un archivo RSS se llama «ítem». Cada ítem consta normalmente de un título, un resumen y un enlace o URL a la página web de origen o que contiene el texto completo. Además, puede contener información adicional como la fecha de publicación o el nombre del autor del texto. El feed RSS se reescribe automáticamente cuando se produce alguna actualización en los contenidos del sitio web. Accediendo al archivo RSS es posible saber si se han actualizado los contenidos y con qué noticias o textos, pero sin necesidad de acceder al sitio web salvo para leer la versión extendida. [32]

## **Ventajas de los Canales RSS:**

1. Da las últimas actualizaciones. Ya sea sobre el tiempo, nueva música, actualización del software, noticias locales, o un nuevo comentario de un sitio de los que raramente se actualizan, pues lo sabe en cuanto sale.
2. Ahorra tiempo al navegar. Desde que el RSS proporciona un resumen del artículo relacionado, ahorra el tiempo del usuario ayudando a decidir qué artículos priorizar cuando lee o busca algo en la red.
3. Disminuye el desorden en su bandeja de entrada. Aunque se le exigirá su dirección de correo electrónico para disfrutar los servicios de los agregators de RSS en línea, RSS no usa su dirección de correo electrónico para enviar las actualizaciones.
4. Está libre de Spam. A diferencia de otras suscripciones de correo electrónico, RSS no hace uso de su dirección para enviar las actualizaciones, porque su dirección se mantiene alejada y segura del Spam.
5. Puede usarse como publicidad, o como una herramienta de Marketing. Los usuarios que subscriben o se sindicán al producto del sitio Web reciben las últimas noticias de los productos y servicios, sin que el sitio Web le envíe correo Spam. Esto es muy ventajoso para el usuario de la red y el dueño del sitio Web, ya que el anuncio se transforma en segmentado; aquéllos que están realmente interesados en sus productos son los que se enteran de los cambios.

## **Desventajas de los Canales RSS:**

1. Los gráficos y fotografías no aparecen en todos los RSS. Para ser más concisos y tener mayor facilidad de publicación, el feed del RSS no despliega las fotografías del sitio original, salvo algún agregador basado en la Web.
2. La identidad del sitio Web puede ser confusa. Desde que la actualización del RSS no despliega el URL real o nombre del sitio Web, puede a veces confundir en lo que un usuario realmente está leyendo.
3. Como es una nueva tecnología, muchos sitios aún no soportan el RSS.

## **Web Scrapping**

Es una técnica utilizada para extraer la información de los sitios web al simular la navegación de un humano por la web. De esta forma se extrae el código HTML y con él, los datos almacenados en la base de datos, permitiendo copiar o duplicar el contenido de un sitio web a otro lugar. Este método ofrece el mejor instrumento posible para recopilar datos, no solo de una sola página web, sino también de todo el dominio, lo que se traduce en la recopilación de una gran cantidad de información de una sola vez. Ofrece un servicio básico que requiere muy poco o prácticamente ningún esfuerzo, y a pesar de su bajo coste estos servicios se encargan de recopilar los datos de las páginas webs de forma eficiente y práctica. En cuanto a la recopilación de datos, errores básicos pueden causar serios problemas, la precisión de los datos extraídos es sumamente importante. Esta técnica permite el monitoreo de las tendencias y últimas noticias en la web. En este sentido, permite la vigilancia de los movimientos de la competencia, así como la opinión de los usuarios sobre el negocio, y de este modo apoya la mejora de procesos para garantizar la mayor satisfacción de los clientes actuales y buscar nuevos.

El scraping es una herramienta muy útil para conocer el comportamiento del mercado. Pero por las características que posee el mismo tiene varias desventajas, una de las más remarcables es en cuanto a las políticas de protección. Pues muchas páginas web tienen prohibido o restringido su uso, esto se debe a que violan restricciones legales y de privacidad al realizar una copia de la información almacenada en el HTML. El proceso de scraping puede resultar confuso para las personas que no tienen experiencia con el mismo debido a la complejidad de la curva de aprendizaje que posee los softwares relacionados. Es común que la activación de las aplicaciones de extracción de datos tarde algún tiempo en iniciarse. Por ello estos servicios pueden tardar varios días hasta que estén completamente operativos y funcionando a pleno rendimiento [32] - [35].

### **Riesgos legales del web scraping:**

Las técnicas de web scraping pueden ser de gran utilidad, pero no siempre están libres de riesgos legales. Puesto que el operador de la web la ha diseñado con usuarios humanos en mente, su apertura automática mediante un web scraper puede suponer un incumplimiento de las condiciones de uso. Estas acciones se vuelven especialmente relevantes cuando se accede a grandes volúmenes de información procedente de varias páginas al mismo tiempo o en sucesión rápida, de un modo en el que una persona nunca sería capaz de interactuar con la página.

Si se realizan de forma automática, la apertura, el almacenamiento y la evaluación de los datos publicados en una página web podrían infringir los derechos de propiedad intelectual. Además, si los datos obtenidos son de carácter personal, almacenarlos y analizarlos sin la autorización de las personas afectadas infringe la normativa vigente de protección de datos. Por esta razón, no está permitido, por ejemplo, scrapear perfiles de Facebook para obtener datos personales [35].

### **Limitaciones técnicas del web scraping:**

Para los operadores de las páginas web suele ser ventajoso limitar las posibilidades de scraping automático en su contenido online. Por un lado, porque el acceso masivo a la web que realizan los scrapers puede perjudicar el rendimiento del sitio y, por otro, porque suele haber secciones internas de la web que no deberían mostrarse en los resultados de búsqueda.

Para limitar el acceso a los scrapers, se ha extendido el uso del estándar robots.txt. Se trata de un archivo de texto que los operadores web ubican en el directorio principal de la página web. En él hay entradas especiales que establecen qué scrapers o bots están autorizados a acceder a qué áreas de la web. Las entradas del archivo *robots.txt* siempre se aplican a un dominio entero [35].

## **1.4. Estado Actual de los Sistemas de Vigilancia Tecnológica en Cuba**

En Cuba desde hace muchos años distintas empresas se han visto interesadas en los procesos de la vigilancia tecnológica, y se ha intentado incursionar en los mismos, la mayoría de los sistemas desarrollados en ese entonces eran simples buscadores, que de una forma u constituyen una etapa de la vigilancia tecnológica.

En los últimos años ya se han realizado avances más concretos en la fundamentación de plataformas desarrolladas específicamente para la vigilancia tecnológica.

### **1.4.1. Plataforma Vigintel**

La empresa de tecnologías de la información (ETI) perteneciente al grupo empresarial de BioCubaFarma, ha desarrollado un sistema de Vigilancia e Inteligencia llamado Vigintel.

El surgimiento del sistema Vigintel ocurrió a partir del conocimiento de la existencia del coronavirus en China, la cual se vio acompañada de las fake news, las cuales se vieron en aumento en las redes sociales en este periodo. Una de las preocupaciones de la comunidad científica en este periodo, además de la búsqueda de tratamientos y vacunas. fue la creación y fortalecimiento de sitios web capaces de brindar la mayor información científica posible sobre la situación epidemiológica en el planeta

El mismo ofrece una variedad de productos informacionales (estudios de mercado, perfiles estratégicos, estudios de tendencias), consultoría para la implementación de un sistema de vigilancia e inteligencia a partir del marco regulador cubano, así como soluciones informáticas desarrolladas bajo los principios de seguridad y sostenibilidad.

También es una plataforma que se utiliza para la automatización del proceso de vigilancia e inteligencia, desde la búsqueda y recuperación de la información hasta la obtención de boletines y alertas. [ 36]

Dentro de sus funcionalidades se encuentran:

- El diseño del despliegue estratégico
- La automatización de la búsqueda y recuperación de información
- La gestión de aportación de valor (validar, categorizar, etiquetar, debatir, colaborar, e idear)
- La creación de alertas y boletines
- Permite gestionar ideas que pueden convertirse en proyectos de investigación alineados con la estrategia, contribuyendo significativamente al proceso de innovación

eti

INICIO NOSOTROS ÁREAS DE NEGOCIO PORTAFOLIO CONTACTO

español Cuba

ación estratégica e innovación: claves para la adopción de la transformación digital...Conocer más | ETI y Cámara de Comercio de Cuba: colaboración a

f t g b

**v** PLATAFORMA VIGINTEL

En el proceso de vigilancia e inteligencia resulta indispensable acceder a un volumen de información localizada en una gran diversidad de fuentes, caracterizadas por su heterogeneidad y dispersión. Esto obliga a los usuarios a dedicar tiempo y esfuerzo a la búsqueda intensiva de este recurso clave en detrimento de fases de análisis, interpretación, difusión y uso del conocimiento generado. **Ante esta situación el uso de herramientas informáticas resulta un imperativo para soportar el ciclo de la vigilancia e inteligencia.** La Empresa de Tecnologías de la Información ofrece la Plataforma VIGINTEL.

**VIGINTEL FACILITA LA AUTOMATIZACIÓN DEL PROCESO DE VIGILANCIA E INTELIGENCIA EN UN ENTORNO COLABORATIVO.**

Vigintel incorpora las políticas de seguridad informática establecidas por la Oficina de Seguridad de Redes Informáticas del Ministerio de Comunicaciones de

Imagen 3. Ilustración de los servicios de Vigintel ofrecidos por ETI [36]

### 1.4.2. DFL Observatorio

Como parte de una estrategia de actualización de su modelo de negocios y aprovechando el marco de la Feria Internacional “Informática 2022”; Dofleini Software presentó DFL Observatorio; una herramienta digital destinada a gestionar la información de una manera clasificada y personalizada, a partir de las especificaciones del usuario.

En este sentido, DFL Observatorio tiene como propósito generar el acceso a diferentes fuentes de información que el cliente seleccione. Estos recursos informativos pueden ser de cualquier género, desde sitios web y revistas científicas, hasta sitios FTP, correos electrónicos y chats de mensajería instantánea. Todo ese contenido es indexado por el software, reconociendo entidades, autores, fechas, orígenes y tipos de documentos, permitiendo la búsqueda organizada y privada de información.

“El producto funciona como un buscador, solo que, en lugar de realizar búsquedas en el universo casi infinito que constituye Internet, lo concentra en aquellas fuentes que el cliente predefinió”, explicó, por su parte, Otto Batista, Especialista Comercial de Dofleini.

“Dichas fuentes son seleccionadas y solicitadas por el usuario a partir de que considera que son verídicas, útiles y de calidad”, agregó. [37]

Ventajas:

- Obtener, clasificar y personalizar la información
- Acceso rápido y ordenado a lo que se busca
- Identificación de expertos sobre las tareas de trabajo
- Generación de alertas sobre nuevos contenidos de interés

Beneficios Empresariales:

- Información segura y centralizada de fácil acceso
- Repositorio de información con integraciones a fuentes de datos externas e internas
- Herramienta para la vigilancia tecnológica y la inteligencia comercial



Imagen 4. DFL Observatorio Tecnológico [ 37]

## 1.5. Tecnologías y Herramientas Utilizadas para dar Solución a la Problemática Propuesta

Se realizan una serie de investigaciones y análisis que permiten comprender el funcionamiento de distintas metodologías y herramientas que servirán para alcanzar el objetivo deseado con el sistema.

### 1.5.1. Metodologías de Software

La metodología hace referencia al conjunto de procedimientos racionales utilizados para alcanzar un objetivo que requiera habilidades y conocimientos específicos. La metodología es una de las etapas específicas de un trabajo o proyecto que parte de una posición teórica y conlleva a una selección de técnicas concretas o métodos acerca del procedimiento para el cumplimiento de los objetivos. Es el conjunto de métodos que se utilizan en una determinada actividad con el fin de formalizarla y optimizarla. Determina los pasos a seguir y cómo realizarlos para finalizar una tarea. [38]

Las metodologías tradicionales son denominadas, a veces, de forma despectiva, como metodologías pesadas. Centran su atención en llevar una documentación exhaustiva de todo el proyecto, la planificación y control del mismo, en especificaciones precisas de requisitos y modelado y en cumplir con un plan de trabajo, definido todo esto, en la fase inicial del desarrollo del proyecto. Estas metodologías tradicionales imponen una disciplina rigurosa de trabajo sobre el proceso de desarrollo del software, se centran especialmente en el control del proceso, mediante una rigurosa definición de roles, actividades, artefactos, herramientas y notaciones para el modelado y documentación detallada. Otra de las características importantes dentro de este enfoque, son los altos costes al implementar un cambio y la falta de flexibilidad en proyectos donde el entorno es volátil.

Por otro lado, se encuentran las metodologías ágiles, este enfoque nace como respuesta a los problemas que puedan ocasionar las metodologías tradicionales y se basa en dos aspectos fundamentales, retrasar las decisiones y la planificación adaptativa. Basan su fundamento en la adaptabilidad de los procesos de desarrollo.

Mediante la siguiente tabla se explican de mejor manera las diferencias existentes entre cada una de las metodologías:

<b>Metodologías Ágiles</b>	<b>Metodologías Tradicionales</b>
Basadas en heurísticas provenientes de prácticas de producción de código	Basadas en normas provenientes de estándares seguidos por el entorno de desarrollo
Especialmente preparados para cambios durante el proyecto	Cierta resistencia a los cambios
Impuestas internamente (por el equipo)	Impuestas externamente
Proceso menos controlado, con pocos principios	Proceso mucho más controlado, con numerosas políticas/normas
No existe contrato tradicional o al menos es bastante flexible	Existe un contrato prefijado
El cliente es parte del equipo de desarrollo	El cliente interactúa con el equipo de desarrollo mediante reuniones
Grupos pequeños (<10 integrantes) y trabajando en el mismo sitio	Grupos grandes y posiblemente distribuidos
Pocos artefactos	Más artefactos
Pocos roles	Más roles
Menos énfasis en la arquitectura del software	La arquitectura del software es esencial y se expresa mediante modelos

Poca documentación	Documentación exhaustiva
Muchos ciclos de entrega	Pocos ciclos de entrega

**Tabla No.1: Comparación entre las metodologías ágiles y tradicional**

Teniendo en cuenta las características presentadas por ambas metodologías y las características del proyecto y del equipo de desarrollo, la metodología a utilizar debe ser una metodología ágil. Entre las principales metodologías ágiles se encuentran las siguientes[39]:

**PROGRAMACIÓN EXTREMA (XP):** La Programación Extrema o Extreme Programming, es un enfoque de la ingeniería de software formulado por Kent Beck, se considera el más destacado de los procesos ágiles de desarrollo de software. Al igual que estos, la programación extrema se diferencia de los métodos tradicionales principalmente en que presenta más énfasis en la adaptabilidad que en la previsibilidad.

**SCRUM:** Según (Orjuela Duarte & Rojas C, 2008), está especialmente indicada para proyectos con un rápido cambio de requisitos. Sus principales características se pueden resumir en dos: El Desarrollo del software se realiza mediante iteraciones denominadas Sprint, con una duración de 30 días, el resultado de cada Sprint es un incremento ejecutable que se muestra al cliente. La Segunda Característica importante son las reuniones a lo largo del proyecto, entre ellas destaca la reunión diaria de 15 minutos del equipo de desarrollo para coordinación e integración.

**CRYSTAL METODOLOGÍAS:** Se Trata de un conjunto de metodologías para el desarrollo de software caracterizadas por encontrarse centradas en las personas que componen equipo y la reducción al máximo del número de artefactos producidos. El desarrollo de software se considera un juego cooperativo de invención y comunicación, limitado por los recursos a utilizar. El equipo de desarrollo es un factor clave, por lo que se deben invertir esfuerzos en mejorar sus habilidades y destrezas, así como tener políticas de trabajo en equipo definidas. Estas políticas dependerán del tamaño del equipo, estableciéndose una clasificación por colores por ejemplo: Crystal Clear (3 a 8 integrantes), Crystal Orange (25 a 50 integrantes).

**ADAPTIVE SOFTWARE DEVELOPMENT (ASD):** Presupone que las necesidades del cliente son cambiantes. La Iniciación de un proyecto involucra definir una misión para él, determinar las características, las fechas y descomponer el proyecto en una serie de pasos individuales, cada uno de los cuales puede abarcar entre cuatro y ocho semanas. Los pasos iniciales deben verificar el alcance del proyecto, los tardíos tienen que ver con el diseño de la arquitectura, la construcción del código, la ejecución de las pruebas finales y el despliegue.

**DYNAMIC SYSTEMS DEVELOPMENT METHOD:** Es la única de las metodologías planteadas surgida de un consorcio formado originalmente por 17 miembros fundadores en enero de 1994. El Objetivo del consorcio era producir una metodología de dominio público que fuera independiente de las herramientas y que pudiera ser utilizada en proyectos de tipo RAD (Rapid Application Development). En DSDM se define cinco fases en la construcción de un sistema, las mismas son: Estudio de factibilidad, Estudio del Negocio, Iteración del modelo funcional, Iteración del diseño y Construcción, Implantación.

### **Metodología de Programación Extrema XP**

La metodología XP es un conjunto de técnicas que dan agilidad y flexibilidad en la gestión de proyectos. También es conocida como Programación Extrema (Extreme Programming) y se centra crear un producto según los requisitos exactos del cliente. De ahí, que le involucre al máximo durante el método de gestión del desarrollo del producto. Es un marco de desarrollo de software ágil que tiene como objetivo producir un software de mayor calidad para mejorar la eficiencia del equipo de desarrollo. Se trata de una metodología de desarrollo cuyo objetivo es promover la aplicación de prácticas de ingeniería apropiadas para la creación de software. Esta metodología la formuló Kent Beck, autor del primer libro sobre este ámbito llamado «Extreme Programming Explained: Embrace Change», publicado en 1999.

Este tipo de programación se diferencia de las metodologías tradicionales en que pone más énfasis en la adaptabilidad que en la previsibilidad. El Extreme Programming considera que los cambios de requisitos sobre la marcha son acciones naturales e inevitables en el desarrollo de un proyecto. Creen que ser capaces de adaptarse a los cambios que puedan surgir en cualquier punto del ciclo de vida de un proyecto es una mejor previsión y más realista que intentarlos definir todos en un principio y que no varíen más [40]

### **Características de la Metodología XP**

- Comunicación constante entre el cliente y el equipo de desarrollo.
- Respuesta rápida a los cambios constantes.
- La planificación es abierta con un cronograma de actividades flexible.
- El software que funciona está por encima de cualquier otra documentación.
- Los requisitos del cliente y el trabajo del equipo del proyecto son los principales factores de éxito del mismo.

### **Herramientas de la Metodología XP**

**Historias de Usuario:** Las Historias de Usuario representan una breve descripción de comportamiento del sistema, se realizan por cada característica principal del sistema y son utilizadas para cumplir estimaciones de tiempo y el plan de lanzamientos, así mismo reemplazan un gran documento de requisitos y presiden la creación de las pruebas de aceptación. Cada historia de usuario debe ser lo suficientemente comprensible y delimitada para que los programadores puedan implementarlas en unas semanas.

**Tareas de Ingeniería:** Una Historia de Usuario se descompone en varias tareas de ingeniería, las cuales describen las actividades que se realizarán en cada historia de usuario, así mismo las tareas de ingeniería se vinculan más al desarrollador, ya que permite tener un acercamiento con el código.

**Pruebas de Aceptación:** Las pruebas de aceptación pertenecen a la categoría de pruebas de caja negra son de vital importancia para el éxito de una iteración y el comienzo de la siguiente, con lo cual el cliente puede conocer el avance en el desarrollo del sistema y a los programadores lo que les resta por hacer. Además, permite una retroalimentación para el desarrollo de las próximas historias de usuarios a ser entregadas. Estas son comúnmente llamadas pruebas del cliente, por lo que son realizadas por el encargado de verificar si las historias de usuarios de cada iteración cumplen con la funcionalidad esperada.

**Las Tarjetas CRC (Clase-Responsabilidades-Colaboradores):** Permiten conocer que clases componen el sistema y cuales interactúan entre sí.

## 1.5.2. Selección de Lenguaje

### Python

Es un lenguaje de programación de alto nivel que se utiliza para desarrollar aplicaciones de todo tipo, creado por Guido Van Rossum a comienzo de los 90. Es un lenguaje de scripting independiente de plataforma y orientado a objetos, preparado para realizar cualquier tipo de programa, desde aplicaciones Windows a servidores de red o incluso, páginas web. A diferencia de otros lenguajes como Java o .NET, Es un lenguaje interpretado, lo que significa que no se necesita compilar el código fuente para poder ejecutarlo, lo que ofrece ventajas como la rapidez de desarrollo e inconvenientes como una menor velocidad. se trata de un lenguaje multiplataforma de código abierto y, por lo tanto, gratuito, lo que permite desarrollar software sin límites, facilita trabajar con inteligencia artificial, big data, machine learning y data science, entre muchos otros campos en auge. [41] [42]

Este constituye un lenguaje muy adecuado para la realización del sistema deseado pues contiene un gran número de librerías que facilitan el trabajo con las API y con el webscraping, sobre todo. Es muy potente y al mismo tiempo es muy sencillo de aprender y de utilizar, además con él se pueden crear aplicaciones ya sean de escritorio o web.



Imagen 5. Lenguaje de programación Python

### 1.5.3. Selección de IDE

#### PyCharm

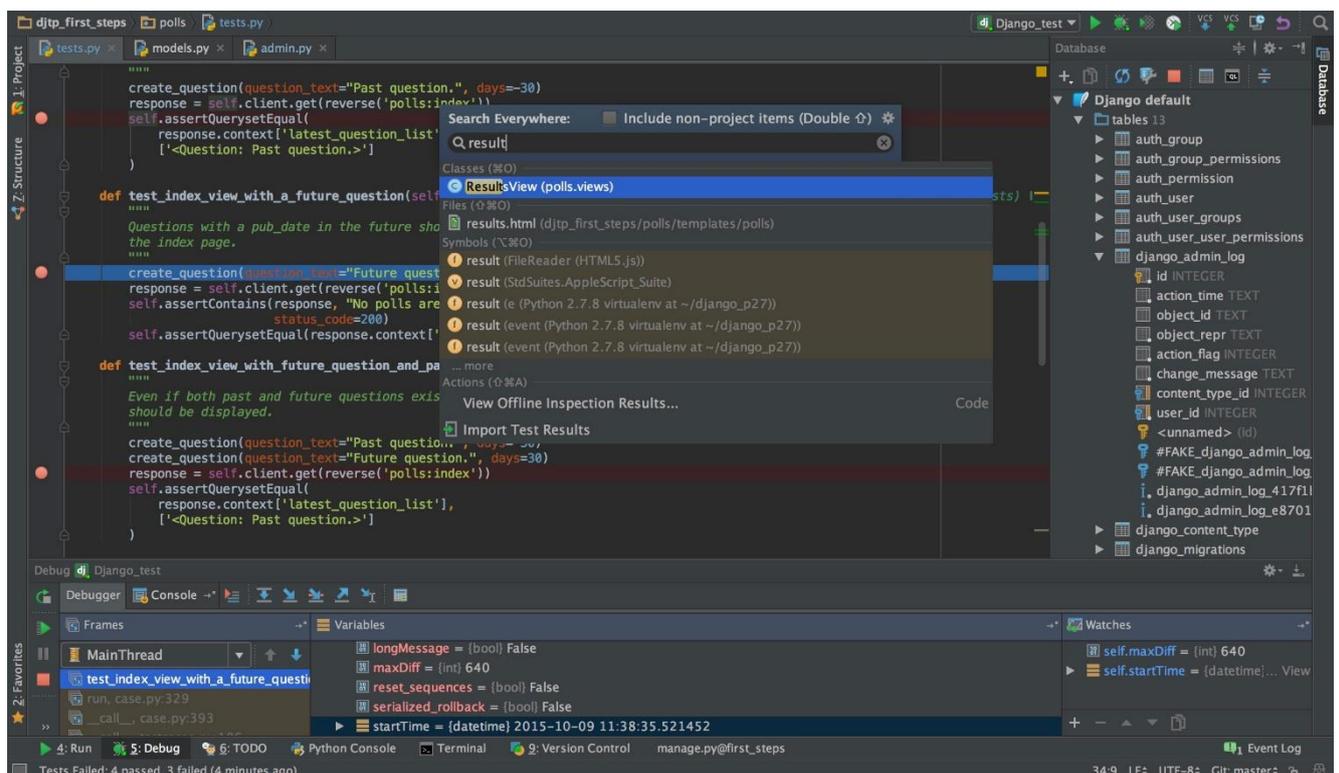
Es un IDE enfocado en Python que trae varias funcionalidades específicamente para este lenguaje. Proporciona una finalización del código inteligente, inspecciones del código, indicación de errores sobre la marcha y arreglos rápidos, así como refactorización de código automática y completas funcionalidades de navegación.

La enorme colección de herramientas pre configuradas de PyCharm incluye un depurador y un ejecutor de pruebas integrados, perfilador Python, un terminal integrado, integración con los principales VCS y herramientas de base de datos integradas, capacidades de desarrollo remoto con intérpretes remotos, un terminal ssh integrado e integración con Docker y Vagrant.

Además de Python, PyCharm ofrece soporte de primer nivel para varios marcos de trabajo de desarrollo web Python, lenguajes de plantilla específicos, JavaScript, CoffeeScript, TypeScript, HTML/CSS, AngularJS, Node.js y más. [43]

#### Compatibilidad Windows [44]

- 64-bit versions of Microsoft Windows 10, 8
- 2 GB RAM minimum, 8 GB RAM recommended
- 2.5 GB hard disk space, SSD recommended
- 1024x768 minimum screen resolution
- Python 2.7, or Python 3.5 or newer



## Imagen 6. PyCharm

### Visual Studio

Es una aplicación informática que proporciona servicios integrales para facilitarle al desarrollador o programador la creación de software, permitiéndonos desarrollar aplicaciones, sitios y aplicaciones web, así como servicios web en cualquier entorno que soporte la plataforma .NET, algunos lenguajes que podemos encontrar son: Visual C# y Visual C++, Java, Python. También se le conoce como un editor de código fuente basado en componentes para crear aplicaciones eficaces y de alto rendimiento, que habilita el uso compartido de herramientas y facilita la creación de soluciones en varios lenguajes, este paquete permite programar contenidos informáticos gráficos de manera simple y accesible, Asimismo, dichos lenguajes utilizan las funciones de .NET Framework, las cuales ofrecen acceso a tecnologías clave para simplificar el desarrollo de aplicaciones web ASP y Servicios Web XML .[45]-[47]

#### Ventajas:

- Es multiplataforma.
- Permite desarrollar con múltiples lenguajes de programación.
- Utiliza la tecnología IntelliSense.
- Utiliza gráficos de vanguardia.

#### Desventajas:

- Problema de versionado asociado con varias librerías runtime DLL, conocido como DLL Hell.
- Soporte pobre para programación orientada a objetos
- Incapacidad para crear aplicaciones multihilo, sin tener que recurrir a llamadas de la API de Windows.
- Dependencia de complejas y frágiles entradas de registro COM

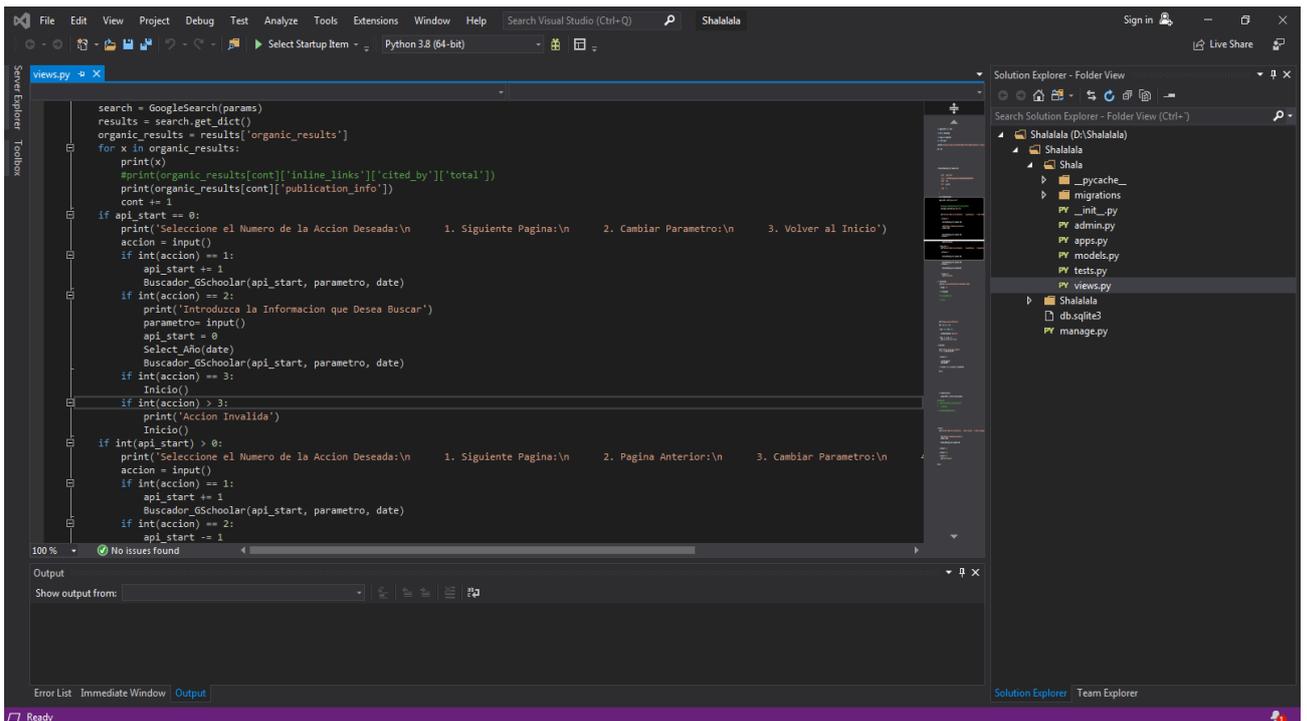


Imagen 7. Visual Studio

A pesar de que el IDE PyCharm sería el ideal para el desarrollo de la plataforma deseada, pues este está especializado para el uso con el lenguaje de programación Python, el mismo cuenta con determinados requisitos para su adecuado funcionamiento, los cuales no son satisfechos en el equipo en el cual se desarrolla la aplicación. Por el motivo anteriormente planteado se utilizará el Visual Studio, el cual, a pesar de no estar hecho específicamente para el trabajo con Python, ofrece herramientas para un trabajo adecuado con el mismo.

#### 1.5.4. Django

Django es un marco framework de Python de alto nivel que fomenta un desarrollo rápido y un diseño limpio y pragmático. Creado por desarrolladores experimentados, se ocupa de gran parte de las molestias del desarrollo web, por lo que puede concentrarse en escribir su aplicación sin necesidad de reinventar la rueda. Es gratis y de código abierto, tiene una comunidad próspera y activa, una gran documentación y muchas opciones de soporte [48]

#### Ventajas:

- Perfecto en función, Características completas: Lleve una gran cantidad de herramientas y marcos de uso común (como la paginación, la autenticación, la gestión de permisos), adecuados para el desarrollo rápido de los sitios empresariales.
- Documento perfecto: después de más de diez años de desarrollo y mejora, Django tiene una amplia gama de casos prácticos y documentos en línea perfectos. Los desarrolladores pueden buscar documentos en línea que buscan soluciones cuando se encuentran con problemas.
- Potentes componentes de acceso a la base de datos: la capa modelo de Django viene con los componentes de la base de datos ORM, lo que hace que los desarrolladores operen en la base de datos sin aprender el idioma SQL.

- Idea de diseño de aplicaciones avanzadas de Django: la aplicación es enchufable, no se piensa mucho. No es necesario, puede eliminarlo directamente, el impacto general en el sistema no es mucho.
- PropioAdministración de la tabla Admin: Solo una plataforma de control de administración de datos de fondo completa solo se puede implementar mediante simples configuraciones y código de varias líneas.
- La información de depuración de Django está agotada: es fácil encontrar el error del código.

### **Desventajas:**

- Gran bolsa: Para algunos de los módulos funcionales que no se requieren para algunas aplicaciones livianas, también se incluye Django, no es tan ligero para el matraz.
- Overpack: Se han encapsulado muchas clases y métodos, y se utilizan directamente, pero es más difícil cambiarlo.
- Desventajas de rendimiento: el rendimiento de Django es bajo en comparación con C, el rendimiento de C ++, por supuesto, esta es la olla de Python, y otros framework de Python tendrán el mismo problema después del tráfico.
- Pregunta de la plantilla: la plantilla de Django implementa el código y el estilo completamente separados, y no permite el código de Python en la plantilla, la flexibilidad puede no ser suficiente para algunos programadores.

## **Conclusiones Parciales**

A lo largo de este capítulo se le dio cumplimiento de manera satisfactoria a los distintos objetivos propuestos. Se realizó la descripción de la Vigilancia Tecnológica, así como la de los Observatorios Tecnológicos, demostrando de esta manera la importancia de ambos, y la necesidad del desarrollo del sistema propuesto para la realización del proceso de Vigilancia Tecnológica en los Observatorios Tecnológicos.

Se determinó como metodología a utilizar la Metodología ágil XP. Fue seleccionado Python como lenguaje de programación y la utilización de su framework Django. Ambos convergerán mediante la utilización del IDE Visual Studio.

# Capítulo 2

## “Propuesta de Solución de la Plataforma Web para la Vigilancia Tecnológica”

### Introducción al Capítulo

El presente capítulo aborda las particularidades del sistema para la búsqueda y recopilación de información a desarrollar. Para lograr obtener las características principales, se hace uso de los requisitos funcionales (RF) y requisitos no funcionales (RnF). Los requisitos describen, las funcionalidades y los atributos de calidad que debe poseer el software. En el capítulo 1, se seleccionó la metodología XP como guía para el desarrollo del software; por lo tanto, se utilizan las HU como herramienta para una descripción detallada de los RF y la confección del plan de iteraciones. Mediante el uso de este último, se procederá a la estimación del tiempo requerido para la culminación del desarrollo del visor y, con el uso de patrones de diseño, se facilita la posterior descripción de las tarjetas CRC.

### 2.1 Descripción de la Propuesta de Solución

Luego de la investigación realizada previamente, se determinó la mejor solución a la situación problemática es la realización de una aplicación web, la cual permitirá a través de la utilización de herramientas como el web scraping, las API y los canales RSS, acceder a la información que se puede obtener mediante estas vías, esto se realizará a través de un vocabulario controlado, una vez obtenidos los resultados esperados, el usuario tendrá la posibilidad de agregarle notas a estos resultados puede asignarle un nivel de valoración, los resultados obtenidos con un número de citas determinados garantizan una valoración de al menos 3, los otros dos puntos los agrega el usuario a dichos resultados. Para llevar a cabo esta aplicación se utilizará la arquitectura cliente servidor y el modelo arquitectónico modelo-vista-templates.

#### 2.1.1 Requisitos Funcionales

Mediante las reuniones con el cliente se lograron identificar 17 requisitos funcionales los cuales se ven plasmados en la siguiente tabla.

**Tabla 1. Requisitos Funcionales**

Nº	Nombre	Prioridad	Complejidad
RF1	Autenticar Usuario	Alta	Media
RF2	Solicitar Información API	Alta	Media
RF3	Listar Información API	Alta	Media
RF4	Filtrar Información API	Baja	Baja
RF5	Agregar Comentario	Alta	Media
RF6	Modificar Comentario	Media	Baja

RF7	Eliminar Comentario	Media	Baja
RF8	Rating Automático	Alta	Media
RF9	Asignar Rating	Alta	Media
RF10	Realizar Scraping	Media	Media
RF11	Solicitar Información RSS	Alta	Media
RF12	Listar Información RSS	Alta	Media
RF13	Filtrar Información RSS	Baja	Baja
RF14	Solicitar Noticias	Alta	Media
RF15	Listar Noticias	Alta	Media
RF16	Ver Detalles	Baja	Baja

## 2.1.2 Requisitos no Funcionales

Requisitos No Funcionales, son requisitos que imponen restricciones en el diseño o la implementación como restricciones en el diseño o Estándares de Calidad. Son propiedades o cualidades que el producto debe tener.

- Usabilidad
  - RNF1. Debe ser una aplicación sustentada en la web
  - RNF2. Debe estar disponible las 24 horas del día
  - RNF3. Está concebida para usuarios con experiencia, pero debe ser intuitiva evitando a si algunos problemas por desconocimiento
- Interfaz Externa
  - RNF4. Debe poseer un diseño sencillo y fácil de usar, pero presentando cierto atractivo visual
- Hardware y software
  - RNF5. El sistema debe ser compatible con las versiones más recientes de Mozilla FireFox y Chrome
- Seguridad
  - RNF6. Se podrá acceder al sistema solamente después de haber sido autenticado con usuario y contraseña
- Diseño e implementación
  - RNF7. El sistema será desarrollado con el framework Django
  - RNF8. Se utilizará Python en su versión 3.8
- Soporte
  - RNF9. El sistema ha sido desarrollado usando Python, por tanto, podrá ser usado en cualquier sistema operativo
- Rendimiento
  - RNF10. El sistema debe responder relativamente rápido a las peticiones del usuario

## 2.2 Descripción de las Historia de Usuario

Tabla 2. Historias de Usuario#1

<b>Numero:</b> 1	<b>Nombre:</b> Autenticar Usuario
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> alta
<b>Puntos estimados:</b> 3	<b>Iteraciones asignadas:</b> 1
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> El sistema debe permitir a los especialistas autenticarse en el sistema mediante la entrada de los datos usuario y contraseñas	
<b>Observaciones:</b> Si a la hora de ingresar los datos algunos de los campos están escrito incorrectamente o le falta algún elemento debe mostrar un mensaje indicando que es lo que ocurre	

Tabla 3. Historias de Usuario#2

<b>Numero:</b> 2	<b>Nombre:</b> Solicitar Información API
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 2	<b>Iteraciones asignadas:</b> 1
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Se selecciona la información a buscar del árbol con el vocabulario controlado y se realiza una llamada las API, en caso de ser un usuario con experiencia puede hacer uso de la barra de búsqueda agilizando el proceso de búsqueda en el árbol	
<b>Observaciones:</b> Si no logra obtener la información, por algún motivo determinado, se debe mostrar el mensaje "No pudo acceder a la información"	

Tabla 4. Historias de Usuario#3

<b>Numero:</b> 3	<b>Nombre:</b> Listar Información API
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 1
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Muestra un listado con la información obtenida por las API	
<b>Observaciones:</b>	

Tabla 5. Historias de Usuario#4

<b>Numero:</b> 4	<b>Nombre:</b> Filtrar Información API
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> baja	<b>Riesgo de desarrollo:</b> baja
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 1
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Los resultados obtenidos mediante las consultas a las API pueden ser filtrados según la fecha.	
<b>Observaciones:</b>	

**Tabla 6. Historias de Usuario#5**

<b>Numero:</b> 5	<b>Nombre:</b> Agregar Comentario
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 2	<b>Iteraciones asignadas:</b> 2
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Cada especialista puede agregar a los resultados obtenidos una serie de comentarios, los cuales podrán ser visualizados por todos los usuarios	
<b>Observaciones:</b>	

**Tabla 7. Historias de Usuario#6**

<b>Numero:</b> 6	<b>Nombre:</b> Modificar Comentario
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> media	<b>Riesgo de desarrollo:</b> baja
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 3
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> El especialista puede modificar cada una de los comentarios adjuntadas a los resultados	
<b>Observaciones:</b>	

**Tabla 8. Historias de Usuario#7**

<b>Numero:</b> 7	<b>Nombre:</b> Eliminar Comentario
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> baja	<b>Riesgo de desarrollo:</b> baja
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 4
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> El especialista puede seleccionar algún comentario publicado por él mismo que desee eliminar	
<b>Observaciones:</b>	

**Tabla 9. Historias de Usuario#8**

<b>Numero:</b> 8	<b>Nombre:</b> Rating Automático
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 2	<b>Iteraciones asignadas:</b> 2
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> La información obtenida según el número de citas que posee se le añade un número determinado de estrellas. El máximo de estrellas que se obtienen de esta manera es 3	
<b>Observaciones:</b>	

**Tabla 10. Historias de Usuario#9**

<b>Numero:</b> 9	<b>Nombre:</b> Asignar Rating
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 2	<b>Iteraciones asignadas:</b> 2
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Cada especialista puede asignarle al resultado el número de estrellas que considere más adecuada y se promediará con el resto de los Rating asignados	
<b>Observaciones:</b>	

**Tabla 11. Historias de Usuario#10**

<b>Numero:</b> 10	<b>Nombre:</b> Realizar Scraping
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> media	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 3
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> El especialista le realiza el scraping a una página determinada mediante la cual no se pueda obtener la información desea mediante las API o canales RSS	
<b>Observaciones:</b> Si no puede realizar el scraping por algún motivo se debe mostrar el mensaje "No se pudo efectuar la Acción deseada"	

**Tabla 12. Historias de Usuario#11**

<b>Numero:</b> 11	<b>Nombre:</b> Solicitar Información RSS
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 2	<b>Iteraciones asignadas:</b> 1
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Se selecciona la información a buscar del árbol con el vocabulario controlado y se realiza una llamada los canales RSS, en caso de ser un usuario con experiencia puede hacer uso de la barra de búsqueda agilizando el proceso de búsqueda en el árbol	
<b>Observaciones:</b> Si no logra obtener la información, por algún motivo determinado, se debe mostrar el mensaje "No pudo acceder a la información"	

**Tabla 13. Historias de Usuario#12**

<b>Numero:</b> 12	<b>Nombre:</b> Listar Información RSS
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 2
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Muestra un listado con la información obtenida por las RSS	
<b>Observaciones:</b>	

**Tabla 14. Historias de Usuario#13**

<b>Numero:</b> 13	<b>Nombre:</b> Filtrar Información RSS
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> baja	<b>Riesgo de desarrollo:</b> baja
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 4
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Los resultados obtenidos mediante las consultas a los canales RSS pueden ser filtrados según la fecha.	

<b>Observaciones:</b>
-----------------------

**Tabla 15. Historias de Usuario#14**

<b>Numero:</b> 14	<b>Nombre:</b> Solicitar Noticia
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 2	<b>Iteraciones asignadas:</b> 1
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Se selecciona la información a buscar del árbol con el vocabulario controlado y que lo que se desea buscar son noticia. Se realiza una llamada a las fuentes de obtención de noticias, en caso de ser un usuario con experiencia puede hacer uso de la barra de búsqueda agilizando el proceso de búsqueda en el árbol	
<b>Observaciones:</b> Si no logra obtener la información, por algún motivo determinado, se debe mostrar el mensaje "No se pudo acceder a la información"	

**Tabla 16. Historias de Usuario#15**

<b>Numero:</b> 15	<b>Nombre:</b> Listar Noticia
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> alta	<b>Riesgo de desarrollo:</b> media
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 2
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Muestra un listado con las Noticias Obtenidas	
<b>Observaciones:</b>	

**Tabla 17. Historias de Usuario#16**

<b>Numero:</b> 16	<b>Nombre:</b> Ver Detalles
<b>Usuario:</b> especialista	
<b>Prioridad en negocio:</b> baja	<b>Riesgo de desarrollo:</b> baja
<b>Puntos estimados:</b> 1	<b>Iteraciones asignadas:</b> 2
<b>Programador Asignado:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Permite al especialista ver más detalles de los resultados que no se muestran en el listado	
<b>Observaciones:</b>	

## 2.3 Plan de Iteraciones

A continuación, se presenta el plan de iteraciones, el cual se encuentra organizado según la importancia dada en las historias de usuario

**Tabla 18. Plan de Iteraciones**

Iteración	HU a implementar	Duración Total(Días)
1	1,2,3,4	10
2	5,6,7,8,9	12
3	10,11,12,13	9
4	14,15,16	8

## 2.4 Plan de Entrega

A continuación, se presenta el plan de entregas correspondiente a la propuesta de solución

**Tabla 19. Plan de Entrega**

Iteración	1	2	3	4
Cantidad de HU	4	5	4	3
Fecha de Entrega	22/10/2022	3/11/2022	12/11/2022	20/11/2022

## 2.5 Estilo Arquitectónico

“La arquitectura de software de un programa o de un sistema computacional está definida por la estructura, comprendida por los elementos de software, las propiedades visibles de esos elementos y las relaciones entre ellos. [49]. Especifica los roles, estructura, organización, reglas, comportamientos, componentes y capas de acuerdo a las características del sistema y las necesidades del cliente, tales restricciones y/o condiciones deben ya estar claramente identificadas en la etapa de levantamiento de requisitos. [50]

Un estilo de arquitectura es un conjunto de decisiones de diseño arquitectural que son aplicables en un contexto de desarrollo específico, restringen las decisiones de diseño de un sistema a ese contexto y plantean como objetivo ciertas cualidades para el sistema resultante. Establecen un vocabulario común, donde se dan nombres a los componentes y conectores, así como a los elementos de datos. Establecen un conjunto de reglas de configuración, como la topología del sistema. Definen una semántica para las reglas de composición de los elementos. Posibilitan el análisis de los sistemas construidos sobre el estilo. [51]

### 2.5.1 Arquitectura Cliente-Servidor

Cliente-Servidor es uno de los estilos arquitectónicos distribuidos más conocidos, el cual está compuesto por dos componentes, el proveedor y el consumidor. El proveedor es un servidor que brinda una serie de servicios o recursos los cuales son consumido por el Cliente. En una arquitectura Cliente-Servidor existe un servidor y múltiples clientes que se conectan al servidor para recuperar todos los recursos necesarios para funcionar, en este sentido, el cliente solo es una capa para representar los datos y se detonan acciones para modificar el estado del servidor, mientras que el servidor es el que realiza todo el trabajo pesado. [52]

Una relación cliente-servidor corresponde al patrón de solicitud-respuesta y debe adherirse al procedimiento de comunicaciones común que define el lenguaje, las reglas o los patrones de diálogo utilizados. La comunicación cliente-servidor se adhiere al conjunto de protocolos TCP o IP. [53]

Mediante esta arquitectura se pueden determinar los siguientes componentes que debe tener la aplicación web a desarrollar:

1. **Lógica de Negocio:**
  - Parte más importante de la aplicación.
  - Define los procesos que involucran a la aplicación.
  - Conjunto de operaciones requeridas para proveer el servicio.
2. **Administración de los datos:**
  - Manipulación de BD y archivos
3. **Interfaz:**
  - Los usuarios acceden a través de navegadores, móviles, PDAs, etc.
  - Funcionalidad accesible a través del navegador.
  - Limitada y dirigida por la aplicación.

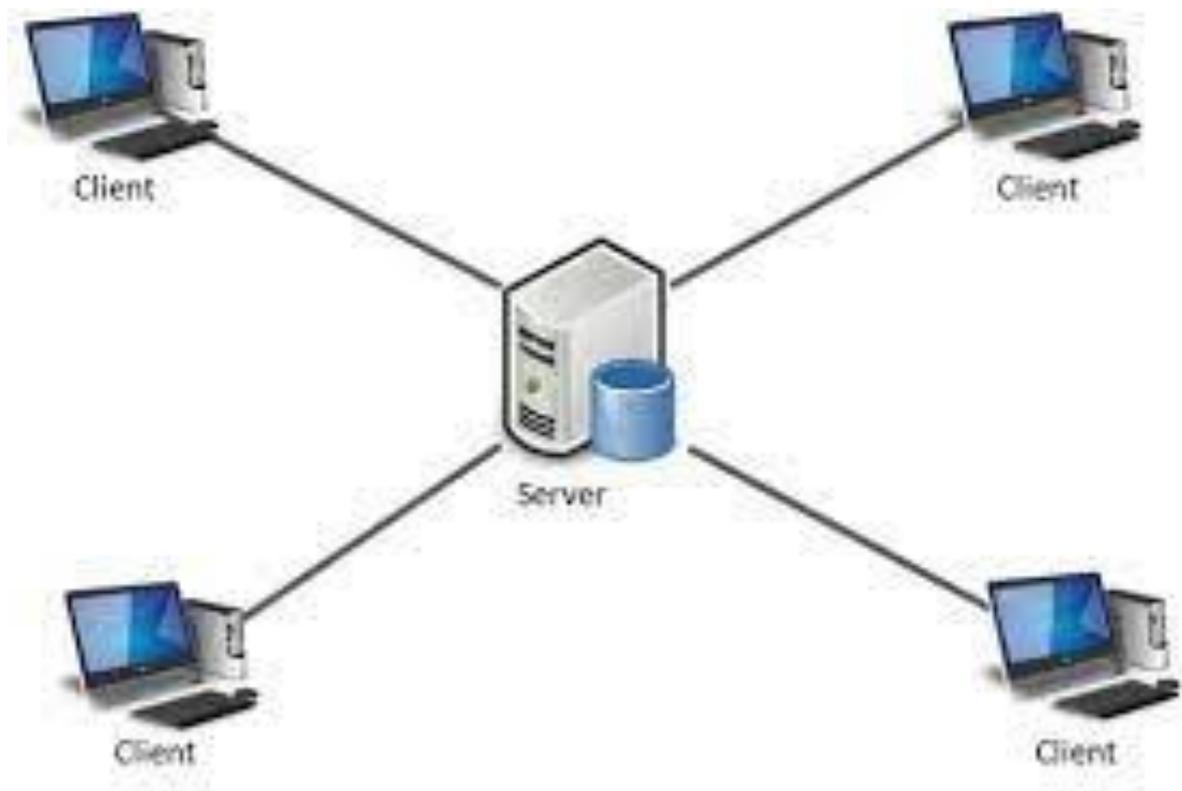


Imagen 8. Representación de la Arquitectura Cliente Servidor

### 2.5.2 Patrón Arquitectónico

El Modelo - Vista -Template, es un patrón de arquitectura de software utilizado por el framework del lenguaje de programación Python, Django. Este como su nombre lo indica está constituido por 3 componentes:

- **Modelo:** la capa de acceso a la base de datos. Esta capa contiene toda la información sobre los datos: cómo acceder a estos, cómo validarlos, cuál es el comportamiento que tiene, y las relaciones entre los datos.
- **Vista:** la capa de la lógica de negocios. Esta capa contiene la lógica que accede al modelo y la delega a la plantilla apropiada: puedes pensar en esto como un puente entre el modelo y las plantillas.
- **Template:** llamada también Plantilla, es la capa de presentación. Esta capa contiene las decisiones relacionadas a la presentación: como algunas cosas son mostradas sobre una página web u otro tipo de documento.

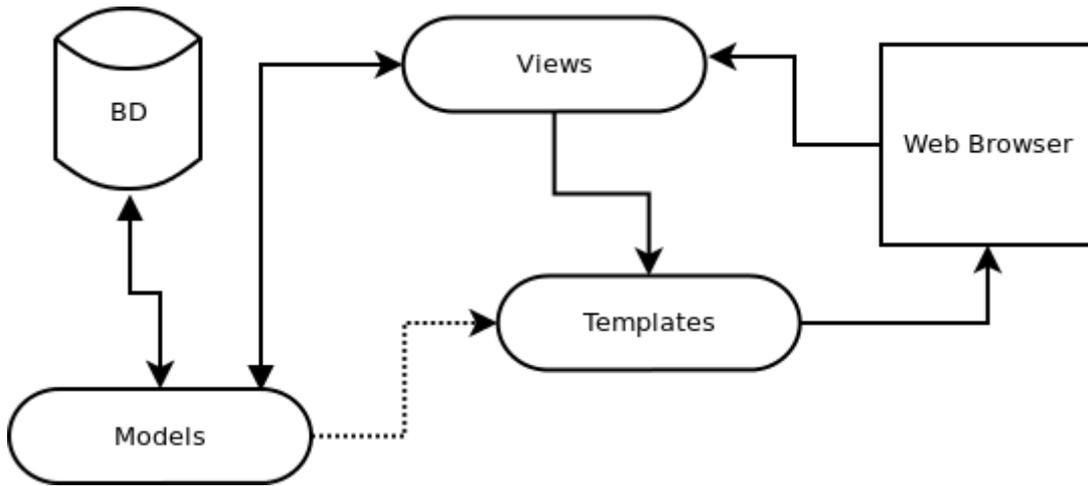


Imagen 9. Representación patrón Modelo-Vista-Template

## 2.6 Tarjetas de Contenido, Responsabilidad y Colaboración

Las tarjetas CRC, presentadas a continuación, se encargan de realizar una representación de los distintos métodos y las relaciones existentes entre cada una de las clases.

Tabla 20. Tarjeta CRC#1

Tarjetas CRC	
Search API	
Responsabilidad	Colaboración
<ul style="list-style-type: none"> <li>• Buscador_GoogleScholar(parametro, cant_result, api_start)</li> <li>• buscar_main()</li> <li>• buscar()</li> <li>• get_apistar()</li> <li>• get_anterior()</li> <li>• get_siguiete()</li> </ul>	Resultado

Tabla 21. Tarjeta CRC#2

Tarjetas CRC	
WebScraping	
Responsabilidad	Colaboración
<ul style="list-style-type: none"> <li>• do_webscraping(url)</li> </ul>	Resultado

Tabla 22. Tarjeta CRC#3

Tarjetas CRC	
CanalesRSS	
Responsabilidad	Colaboración

<ul style="list-style-type: none"> <li>• buscar_rss(parametro, cant_result, start)</li> <li>• buscar_main()</li> <li>• buscar()</li> <li>• get_start()</li> <li>• get_anterior()</li> <li>• get_siguiete()</li> </ul>	Resultado Search API
---	----------------------

**Tabla 23. Tarjeta CRC#4**

Tarjetas CRC	
Search Noticias	
Responsabilidad	Colaboración
<ul style="list-style-type: none"> <li>• buscar_noticias(parametro, api_start)</li> <li>• buscar_main()</li> <li>• buscar()</li> <li>• get_apistar()</li> <li>• get_anterior()</li> <li>• get_siguiete()</li> </ul>	Resultado Noticias

## Conclusiones Parciales

A lo largo de este capítulo se describieron las características generales del sistema para la búsqueda y recopilación de información. Se identificaron un total de 16 RF y 10 RnF. Mediante el uso de los RF, se redactaron las HU, permitiendo posteriormente, la creación de un plan de 4 iteraciones y una estimación de 5.5 semanas para desarrollar el sistema. La utilización del framework django facilitó la selección del patrón de diseño, siendo este Modelo Vista Template. Una vez determinado el patrón se pudo realizar una descripción adecuada de las tarjetas CRC.

# Capítulo 3

## “Implementación y pruebas”

### Introducción al Capítulo

El desarrollo del este capítulo se centrará en la descripción de la etapa de implementación y las distintas pruebas de software. De este modo se inicia con las tareas de ingeniería, las cuales se realizan teniendo en cuenta el punto de vista del programador. Posteriormente para garantizar una mayor comprensión de la plataforma se realiza la descripción de la misma, así como su composición visual. Con la culminación de esta última etapa se procede con la realización de las pruebas de validación, para garantizar el adecuado funcionamiento de la página.

### 3.1 Tareas de Ingeniería

Las tareas de ingeniería parten de la descomposición de las HU, estas describen las actividades que se realizarán en cada HU. Cada HU puede ser descompuesta en distintas tareas de ingeniería. Estas tareas se vinculan más al desarrollador, pues permite tener un acercamiento con el código y describes los pasos necesarios para solucionar las necesidades planteadas en las HU.

### 3.1.1 Iteración I

En la primera iteración se encuentran las funcionalidades relacionadas con la realización de las peticiones para búsqueda y la visualización de los resultados de los resultados obtenidos, así como la autenticación del usuario: Las tareas definidas son:

- Tarea No.1: Validar Credenciales
- Tarea No.2: Autenticar
- Tarea No.3: Realizar petición API partiendo del TreeView
- Tarea No.4: Realizar petición API partiendo de la Barra de Búsqueda
- Tarea No.5: Cargar Resultados API
- Tarea No.6: Filtrar Resultados API

Tabla 24. Tarea#1

Tarea	
Número de tarea:1	Historia de Usuario:1
Nombre de la tarea: Validar Credenciales	
Tipo de Tarea: Desarrollo	Puntos estimados:0.5
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que sea capaz verificar que los datos introducidos por el usuario cumplen con los requisitos necesarios	

Tabla 25. Tarea#2

Tarea	
Número de tarea:2	Historia de Usuario:1
Nombre de la tarea: Autenticar	
Tipo de Tarea: Desarrollo	Puntos estimados:1
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que sea capaz de determinar si el usuario ingresado y su contraseña se encuentran registrados en la base de datos	

Tabla 26. Tarea#3

Tarea	
Número de tarea:3	Historia de Usuario:2
Nombre de la tarea: Realizar petición API partiendo del TreeView	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que obtenga el valor de la hoja del árbol deseada y se realice la búsqueda de la información relacionada	

Tabla 27. Tarea#4

Tarea	
Número de tarea:4	Historia de Usuario:2
Nombre de la tarea: Realizar petición API partiendo de la Barra de Búsqueda	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción Crear un método que obtenga el valor introducido en el input y se realice la búsqueda de la información relacionada	

Tabla 28. Tarea#5

Tarea	
Número de tarea:5	Historia de Usuario:3
Nombre de la tarea: Cargar Resultados API	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que sea capaz realizar las búsquedas solicitadas partiendo de los datos introducidos o seleccionados	

Tabla 29. Tarea#6

Tarea	
Número de tarea:6	Historia de Usuario:4
Nombre de la tarea: Filtrar Resultados API	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que sea capaz filtrar los resultados obtenidos teniendo en cuenta la fecha de publicación del artículo o el idioma	

### 3.1.2 Iteración II

En la segunda iteración se encuentran las funcionalidades relacionadas con la asignación de comentarios y la asignación de rating a los resultados obtenidos. Las tareas definidas son:

- Tarea No.7: Agregar Comentario
- Tarea No.8: Modificar Comentario
- Tarea No.9: Eliminar Comentario
- Tarea No.10: Rating Automático
- Tarea No.11: Asignar Rating

Tabla 30. Tarea#7

Tarea	
Número de tarea:7	Historia de Usuario:5
Nombre de la tarea: Agregar Comentario	

<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.2
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que permita a los usuarios agregar comentarios sobre algún resultado de búsqueda en específico	

Tabla 31. Tarea#8

<b>Tarea</b>	
<b>Número de tarea:</b> 8	<b>Historia de Usuario:</b> 6
<b>Nombre de la tarea:</b> Modificar comentario	
<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.2
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que permita a los usuarios modificar las entradas que ellos mismos hayan creado	

Tabla 32. Tarea#9

<b>Tarea</b>	
<b>Número de tarea:</b> 9	<b>Historia de Usuario:</b> 7
<b>Nombre de la tarea:</b> Eliminar Comentario	
<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.2
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que permita al usuario eliminar un comentario realizado sobre un resultado determinado	

Tabla 33. Tarea#10

<b>Tarea</b>	
<b>Número de tarea:</b> 10	<b>Historia de Usuario:</b> 8
<b>Nombre de la tarea:</b> Asignar rating automático	
<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.4
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que agregue un nivel de rating o relevancia automáticamente según la cantidad de citas que presente	

Tabla 34. Tarea#11

<b>Tarea</b>	
<b>Número de tarea:</b> 11	<b>Historia de Usuario:</b> 9
<b>Nombre de la tarea:</b> Asignar rating	
<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.4
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que le permita al usuario agregar la valoración o nivel de relevancia de un artículo según su criterio propio	

### 3.1.3 Iteración III

En la tercera iteración se encuentran las funcionalidades relacionadas con modificar los comentarios, realizar el web scraping, y realizar las búsquedas utilizando los canales RSS. Las tareas definidas son:

- Tarea No.12: Realizar el Scrapping
- Tarea No.13: Realizar petición RSS partiendo del TreeView
- Tarea No.14 Realizar petición RSS partiendo de la Barra de Búsqueda
- Tarea No.15: Cargar Resultados RSS
- Tarea No.16: Filtrar Resultados RSS

Tabla 35. Tarea#12

Tarea	
Número de tarea:12	Historia de Usuario:10
Nombre de la tarea: Realizar el Scrapping	
Tipo de Tarea: Desarrollo	Puntos estimados:0.5
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que permita la realización del web scraping a la url obtenida por algún resultado deseado	

Tabla 36. Tarea#13

Tarea	
Número de tarea:13	Historia de Usuario:11
Nombre de la tarea: Realizar petición RSS partiendo del TreeView	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que obtenga el valor de la hoja del árbol deseada y se realice la búsqueda de la información relacionada	

Tabla 37. Tarea#14

Tarea	
Número de tarea:14	Historia de Usuario:11
Nombre de la tarea: Realizar petición RSS partiendo de la Barra de Búsqueda	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción Crear un método que obtenga el valor introducido en el input y se realice la búsqueda de la información relacionada	

Tabla 38. Tarea#15

Tarea	
Número de tarea:15	Historia de Usuario:12
Nombre de la tarea: Cargar Resultados RSS	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:

<b>Programador responsable:</b> Luyser Hernández Fariñas
<b>Descripción:</b> Crear un método que sea capaz realizar las búsquedas solicitadas partiendo de los datos introducidos o seleccionados

Tabla 39. Tarea#16

<b>Tarea</b>	
<b>Número de tarea:</b> 16	<b>Historia de Usuario:</b> 13
<b>Nombre de la tarea:</b> Filtrar Resultados RSS	
<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.4
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que sea capaz filtrar los resultados obtenidos teniendo en cuenta la fecha de publicación del artículo o el idioma	

### 3.1.4 Iteración IV

En la cuarta iteración se encuentran las funcionalidades relacionadas con las búsqueda y obtención de noticias, así como la función que permite ver más detalles de los resultados obtenidos:

- Tarea No.17: Realizar petición Noticias partiendo del TreeView
- Tarea No.18: Realizar petición Noticias partiendo de la Barra de Búsqueda
- Tarea No.19: Cargar Noticias
- Tarea No.20: Ver Detalles

Tabla 40. Tarea#17

<b>Tarea</b>	
<b>Número de tarea:</b> 17	<b>Historia de Usuario:</b> 14
<b>Nombre de la tarea:</b> Realizar petición Noticias partiendo del TreeView	
<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.4
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que obtenga el valor de la hoja del árbol deseada y se realice la búsqueda de la información relacionada	

Tabla 41. Tarea#18

<b>Tarea</b>	
<b>Número de tarea:</b> 18	<b>Historia de Usuario:</b> 14
<b>Nombre de la tarea:</b> Realizar petición Noticias partiendo de la Barra de Búsqueda	
<b>Tipo de Tarea:</b> Desarrollo	<b>Puntos estimados:</b> 0.4
<b>Fecha de inicio:</b>	<b>Fecha de fin:</b>
<b>Programador responsable:</b> Luyser Hernández Fariñas	
<b>Descripción:</b> Crear un método que obtenga el valor introducido en el input y se realice la búsqueda de la información relacionada	

Tabla 42. Tarea#19

Tarea	
Número de tarea:19	Historia de Usuario:15
Nombre de la tarea: Listar Noticias	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que sea capaz de mostrar en pantalla las información solicitada	

Tabla 43. Tarea#20

Tarea	
Número de tarea:20	Historia de Usuario:16
Nombre de la tarea: Visualizar Detalles	
Tipo de Tarea: Desarrollo	Puntos estimados:0.4
Fecha de inicio:	Fecha de fin:
Programador responsable: Luyser Hernández Fariñas	
Descripción: Crear un método que sea capaz mostrar toda la información posible obtenida de los resultados de las búsquedas realizadas	

## 3.2 Descripción y Funcionamiento de la Plataforma

### 3.2.1 Página Principal

Para la realización de una plataforma para la vigilancia tecnológica, hay que tener definido el tipo de información que se desea monitorear. En este caso se monitoreará toda la información posible relacionada con las ciencias de la computación. Tratando de abarcar un número considerable de temas relacionados las misma. Esto se logrará utilizando un abanico de términos derivados de dicha ciencia. Como se expuso en capítulos anteriores para poder realizar las peticiones se recurrirá a un vocabulario controlado el cual favorece a garantizar una mayor calidad en los resultados obtenidos.

Al ingresar a la página principal del sitio se podrá observar a la izquierda de la pantalla ocupando la mayor parte de la misma un árbol. Cada una de las ramas de dicho árbol representa temas y subtemas derivados de las ciencias de la computación cada rama se puede ir despegando en más y más ramas hasta llegar a las hojas las cuales contienen palabras claves que ayudaran a realizar la búsqueda. El objetivo de esta distribución en forma de árbol es ayudar a los usuarios o poder realizar una búsqueda de calidad pues funciona de ayuda y guía para alcanzar la información deseada. La utilización de este árbol es recomendable para usuarios que recién comienzan a utilizar la aplicación, o como se planteaba anteriormente en caso de no estar la idea concreta de que desea buscar este le ayudara a dar con la respuesta deseada.

A la derecha de la pantalla se muestra una pequeña barra de búsqueda. El objetivo de esta es servir de apoyo al árbol, para los usuarios que tengan experiencia y ya sepan que es lo desean buscar en concreto. Esto facilita el hecho de tener que dar un gran

número de clicks para acceder a la información deseada. Una vez conocida las distintas palabras claves el usuario puede acceder directamente a ella escribiéndola en el buscador. Este buscador cuenta a su vez con una herramienta que le es común tanto a árbol como a la barra, es un campo desplegable el cual le permite al usuario seleccionar el tipo de documentación que desea. Le permite determinar si lo que está buscando son artículos científicos, código, o noticias.

### **3.2.2 Página Resultado**

Una vez realizada la petición en la página principal se carga la página con los resultados obtenidos la misma en el lado derecho de la pantalla un listado con los resultados. De los mismos se muestra a primera vista el título y un breve resumen. En la esquina inferior izquierda de cada uno de los resultados se pueden ver dos botones uno de ellos permite ver todos los detalles posibles obtenidos, se vuelve a mostrar el título y el resumen, pero esta vez acompañado del nombre de los autores, así como el link y la cantidad de citas que posee. Esta opción también muestra un botón que le permite al usuario hacer una llamada al método web scrapping el cual devuelve el código html del sitio deseado sin necesidad de acceder directamente al mismo, esto permite analizar más a fondo los resultados obtenidos, también le da la posibilidad al usuario, si la página fue de gran interés y utilidad guardar el resultado del scrapping realizado.

El botón que se encuentra junto al que permite ver los detalles del resultado, permite agregarle un comentario a la búsqueda deseada, de este modo cada uno de los usuarios puede dejar su opinión sobre lo que pudo observar. De este modo todos los usuarios pueden mantener una retroalimentación sobre que artículos pueden ser de utilidad y cuales valdría la pena realizarle un seguimiento más especializado.

En la parte superior del contenedor donde se muestran los resultados se pueden apreciar dos botones y un campo de texto en el medio. Estos permiten a los usuarios moverse mediante las distintas paginas obtenidas de los resultados. Los botones que indican siguiente y anterior permiten al usuario como bien dice su nombre desplazarse hacia página siguiente o la anterior, el input que se haya entre los dos botones, al igual que estos últimos permite al usuario desplazarse entre cada una de las paginas, pero a diferencia de los anteriores este re direcciona al usuario directamente a la página que se ingrese.

Ocupando una pequeña parte del sector derecho de la pantalla se encuentra un pequeño menú el cual contiene un input, una lista desplegable y varios elementos más. Estos elementos cumplen con la función de personalizar la búsqueda realizada, ya sea cambiar el número de resultados mostrados en pantalla, el o los idiomas en los que se encuentran los resultados y el rango de tiempo en el que se quiere obtener los datos. Y al igual que en la principal el usuario con experiencia puede realizar desde la misma página de los resultados una nueva búsqueda seleccionando el tipo de información que desea obtener (si es noticia, o artículo) esto último mediante una lista desplegable.

### 3.3 Implementación

El sistema desarrollado, para asegurar el cumplimiento de los requerimientos planteados, fue necesario la utilización de diversas librerías. De entre todas las librerías existentes en el proyecto existen tres, de las cuales posteriormente se derivan los métodos que se consideran los más importantes para la búsqueda y recopilación de información. Estas librerías son: **serapi** y **feedparser**.

En el caso de serapi esta es la encargada de posibilitar la comunicación entre el sistema desarrollado y las distintas API ofrecidas por Google. Mediante el uso de la misma se puede acceder a fuentes de artículos científicos como es el caso de GoogleScholar, y fuentes de noticias como es el caso de GoogleNews. A continuación, se muestra un fragmento del código donde se realiza el llamado a la API de GoogleScholar.

#### Código Fuente 3.1. Obtención de resultados mediante Google Scholar

```
def buscador_gschoolar(parametro, cant_result, api_start):  
  
    ListAtributos=[]  
    params = {  
        'q'      : parametro,  
        'num'    : cant_result,  
        'start'  : api_start,  
        'engine' : "google_scholar",  
        'api_key' : '48777cc34a4298ecd1dde173d0c1aab9e71e40b70a70be263f407e5691cb5517',  
        'output' : 'html',  
        'lr'     : 'lang_en|lang_es',  
        'as_vis' : '0',  
        'scisbd' : '0',  
    }  
  
    search = GoogleSearch(params)  
    results = search.get_dict()  
    organic_results = results['organic_results']
```

La librería feedparser es la encargada de realizar la solicitud y el procesamiento de los archivos XML obtenidos mediante los canales RSS. Utilizando esta librería, con tan solo realizarle la petición utilizando la dirección web del canal RSS que se desea consumir, este parsea el archivo obtenido posibilitando el consumo de los datos deseados presentes en el mismo. A continuación se muestra el método mediante el cual se realiza la petición y se recorre la misma realizando las peticiones deseadas.

### Codigo Fuente 3.2. Obtencion de resultados mediante RSS

```
def buscar_rss(parametro, cant_result, start, ListAtributos):
    x= start+1
    start= str(x)
    url =
    "https://www.lareferencia.info/vufind/Search/Results?sort=last_indexed+desc&limit={0}&view=rss&page={1}&lookfor={2}&type=AllFields"
    rss= feedparser.parse(url.format(cant_result, start , parametro.replace(" ", "+")))
    for post in rss.entries:
        Atributos= Resultado()
        Atributos.titulo=post.title
        Atributos.link=post.guid
        Atributos.id=post.link
        Atributos.citas="Desconocido"
        Atributos.resumen=""
        Atributos.autor=post.author
        ListAtributos.append(Atributos)
```

## 3.4 Pruebas de Aceptación

En capítulos anteriores cuando se determinaron las herramientas y tecnologías a utilizar; se pudo apreciar que la metodología XP propone las pruebas de aceptación, como vía para asegurar que se cumplan de manera efectiva los requisitos planteados por el cliente. Las mismas se realizan en cada una de las iteraciones, por este motivo es necesario la obtención de buenos resultados para poder avanzar hacia la siguiente iteración. A continuación se muestran las pruebas realizadas.

### 3.4.1 Iteración I

Tabla 44. Prueba de aceptación#1

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU1_P1	<b>Historia de Usuario:</b> 1
<b>Nombre:</b> Validar Credenciales	
<b>Descripción:</b> Prueba para comprobar que los datos introducidos son correctos	
<b>Condiciones de ejecución:</b> El usuario debe introducir sus credenciales	
<b>Pasos de ejecución:</b> El usuario debe escribir su nombre de usuario y su contraseña.	

<b>Resultados esperados:</b> De ser correctos los datos se procede a comprobar que el usuario exista
--

**Tabla 45. Prueba de aceptación#2**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU1_P2	<b>Historia de Usuario:</b> 1
<b>Nombre:</b> Validar Credenciales	
<b>Descripción:</b> Prueba para comprobar que los datos introducidos son correctos	
<b>Condiciones de ejecución:</b> El usuario debe introducir sus credenciales	
<b>Pasos de ejecución:</b> El usuario debe escribir su nombre de usuario y su contraseña.	
<b>Resultados esperados:</b> De ser incorrectos los datos, se muestra un mensaje "Existen campos incorrectos"	

**Tabla 46. Prueba de aceptación#3**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU2_P1	<b>Historia de Usuario:</b> 2
<b>Nombre:</b> Autenticar	
<b>Descripción:</b> Comprueba que el usuario introducido sea un usuario válido	
<b>Condiciones de ejecución:</b> El usuario debe introducir sus credenciales	
<b>Pasos de ejecución:</b> El usuario debe escribir su nombre de usuario y su contraseña.	
<b>Resultados esperados:</b> Si los datos introducidos por el usuario coinciden con la base de datos de usuarios se accede a la página principal	

**Tabla 47. Prueba de aceptación#4**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU2_P2	<b>Historia de Usuario:</b> 2
<b>Nombre:</b> Autenticar	
<b>Descripción:</b> Comprueba que el usuario introducido sea un usuario válido	
<b>Condiciones de ejecución:</b> El usuario debe introducir sus credenciales	
<b>Pasos de ejecución:</b> El usuario debe escribir su nombre de usuario y su contraseña.	
<b>Resultados esperados:</b> De ser incorrectos los datos, se muestra el mensaje "El usuario no existe"	

**Tabla 48. Prueba de aceptación#8**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU3_P1	<b>Historia de Usuario:</b> 3
<b>Nombre:</b> Realizar petición API partiendo del TreeView	
<b>Descripción:</b> Prueba para asegurarse que se puedan escoger los datos a buscar desde el árbol	
<b>Condiciones de ejecución:</b> El usuario debe seleccionar la hoja que represente la información que se desea	
<b>Pasos de ejecución:</b> El usuario debe desplegar las ramas del árbol que le interesan hasta llegar a las hojas y seleccionar la hoja con la información que se desea buscar	
<b>Resultados esperados:</b> Una vez seleccionada la información deseada se envía al backend y se procede a realizar la búsqueda de los resultados	

**Tabla 49. Prueba de aceptación#6**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU4_P1	<b>Historia de Usuario:</b> 4
<b>Nombre:</b> Realizar petición API partiendo del Barra de Búsqueda	
<b>Descripción:</b> Prueba para asegurarse que se pueda realizar la petición partiendo desde la barra de búsqueda	
<b>Condiciones de ejecución:</b> El usuario introduce el dato deseado	

<b>Pasos de ejecución:</b> El usuario debe introducir en el campo de la barra de búsqueda el parámetro deseado para posteriormente realizar la búsqueda
<b>Resultados esperados:</b> Una vez introducida la información que se desea buscar, de coincidir con las hojas existentes en el árbol, se realiza la búsqueda de la información

**Tabla 50. Prueba de aceptación#7**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU4_P2	<b>Historia de Usuario:</b> 4
<b>Nombre:</b> Realizar petición API partiendo del Barra de Búsqueda	
<b>Descripción:</b> Prueba para asegurarse que se pueda realizar la petición partiendo desde la barra de búsqueda	
<b>Condiciones de ejecución:</b> El usuario introduce el dato deseado	
<b>Pasos de ejecución:</b> El usuario debe introducir en el campo de la barra de búsqueda el parámetro deseado para posteriormente realizar la búsqueda	
<b>Resultados esperados:</b> Una vez introducida la información que se desea buscar, si no coincide con las hojas del árbol muestra el mensaje “Los parámetros no son correctos”	

**Tabla 51. Prueba de aceptación#8**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU5_P1	<b>Historia de Usuario:</b> 5
<b>Nombre:</b> Cargar Resultados API	
<b>Descripción:</b> Se realizan las búsquedas de la información deseada	
<b>Condiciones de ejecución:</b> Cargar el listado que muestra los resultados obtenidos	
<b>Pasos de ejecución:</b> Una vez introducido los datos seleccionados se realizan las llamadas a las API, para la búsqueda de la información	
<b>Resultados esperados:</b> De ser satisfactoria la búsqueda se muestra el listado con los resultados obtenidos	

**Tabla 52. Prueba de aceptación#9**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU5_P2	<b>Historia de Usuario:</b> 5
<b>Nombre:</b> Cargar Resultados API	
<b>Descripción:</b> Se realiza la búsqueda de la información deseada	
<b>Condiciones de ejecución:</b> Cargar el listado que muestra los resultados obtenidos	
<b>Pasos de ejecución:</b> Una vez introducido los datos seleccionados se realizan la llama a la API, para la búsqueda de la información	
<b>Resultados esperados:</b> Muestra el mensaje “No se pudo obtener resultados”	

**Tabla 53. Prueba de aceptación#10**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU6_P1	<b>Historia de Usuario:</b> 6
<b>Nombre:</b> Filtrar Resultados API	
<b>Descripción:</b> Se seleccionan los datos por los que se desea filtrar la información	
<b>Condiciones de ejecución:</b> Seleccionar los atributos de filtrado	
<b>Pasos de ejecución:</b> El usuario selecciona los atributos de filtrado deseados, como la cantidad de resultados a mostrar en pantalla y la rango de años de los mismos	
<b>Resultados esperados:</b> Se cargan los resultados cumpliendo con los atributos de filtrado seleccionados	

### 3.4.2 Iteración II

Tabla 54. Prueba de aceptación#11

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU7_P1	<b>Historia de Usuario:</b> 7
<b>Nombre:</b> Agregar Comentario	
<b>Descripción:</b> El usuario selecciona el resultado al que le desea agregar un comentario	
<b>Condiciones de ejecución:</b> Se debe seleccionar el resultado al que agregarle comentario	
<b>Pasos de ejecución:</b> El usuario se dirige y presiona en el botón comentario y una vez dentro puede escribir el comentario deseado	
<b>Resultados esperados:</b> Se agrega a la lista el comentario introducido por el usuario	

Tabla 55. Prueba de aceptación#12

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU8_P1	<b>Historia de Usuario:</b> 8
<b>Nombre:</b> Modificar Comentario	
<b>Descripción:</b> El usuario selecciona modifica el comentario seleccionado	
<b>Condiciones de ejecución:</b> El comentario debe haber sido escrito por el usuario	
<b>Pasos de ejecución:</b> Se selecciona el comentario que se desea modificar y luego se dirige al botón editar y le permite cambiar el comentario seleccionado	
<b>Resultados esperados:</b> Se modifica satisfactoriamente el comentario seleccionado	

Tabla 56. Prueba de aceptación#13

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU9_P1	<b>Historia de Usuario:</b> 9
<b>Nombre:</b> Agregar Comentario	
<b>Descripción:</b> El usuario selecciona el comentario que desea eliminar	
<b>Condiciones de ejecución:</b> El comentario debe haber sido escrito por el usuario	
<b>Pasos de ejecución:</b> Se selecciona el comentario que se desea eliminar y luego se dirige al botón eliminar y le permite borrar el comentario seleccionado	
<b>Resultados esperados:</b> Se elimina satisfactoriamente el comentario seleccionado	

Tabla 57. Prueba de aceptación#14

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU10_P1	<b>Historia de Usuario:</b> 10
<b>Nombre:</b> Rating Automático	
<b>Descripción:</b> Cada resultado obtenido muestra un número de estrellas indicando el rating	
<b>Condiciones de ejecución:</b> Debe haberse realizado una búsqueda previamente	
<b>Pasos de ejecución:</b> El resultado obtenido muestra un número de estrellas de manera automática como máximo tres en correspondencia al numero de citas que posea en caso de no poseer citas o ser desconocida se le considera como 1 estrella	
<b>Resultados esperados:</b> Se muestran las estrellas en pantalla por cada resultado	

Tabla 58. Prueba de aceptación#15

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU11_P1	<b>Historia de Usuario:</b> 11
<b>Nombre:</b> Asignar Rating	
<b>Descripción:</b> Cada resultado obtenido muestra un número de estrellas indicando el rating	
<b>Condiciones de ejecución:</b> Debe haberse realizado una búsqueda previamente	

<b>Pasos de ejecución:</b> El usuario es capaz de modificar el rating añadiendo el que se considere pertinente, posteriormente se realiza un cálculo del promedio del rating asociado al resultado y se muestra en la estrella
<b>Resultados esperados:</b> Se muestran las estrellas en pantalla por cada resultado

### 3.4.3 Iteración III

**Tabla 59. Prueba de aceptación#16**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU12_P1	<b>Historia de Usuario:</b> 12
<b>Nombre:</b> Rating Automático	
<b>Descripción:</b> El usuario selecciona el resultado al que se le va a realizar el webscrapping	
<b>Condiciones de ejecución:</b> Debe haberse realizado una búsqueda previamente	
<b>Pasos de ejecución:</b> El usuario selecciona el resultado de búsqueda al que desea realizarle el webscrapping	
<b>Resultados esperados:</b> Se muestra en pantalla los resultados obtenidos mediante el webscrapping	

**Tabla 60. Prueba de aceptación#17**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU13_P1	<b>Historia de Usuario:</b> 13
<b>Nombre:</b> Realizar petición RSS partiendo del TreeView	
<b>Descripción:</b> Prueba para asegurarse que se puedan escoger los datos a buscar desde el árbol	
<b>Condiciones de ejecución:</b> El usuario debe seleccionar la hoja que represente la información que se desea	
<b>Pasos de ejecución:</b> El usuario debe desplegar las ramas del árbol que le interesan hasta llegar a las hojas y seleccionar la hoja con la información que se desea buscar	
<b>Resultados esperados:</b> Una vez seleccionada la información deseada se envía al backend y se procede a realizar la búsqueda de los resultados	

**Tabla 61. Prueba de aceptación#18**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU14_P1	<b>Historia de Usuario:</b> 14
<b>Nombre:</b> Realizar petición RSS partiendo del Barra de Búsqueda	
<b>Descripción:</b> Prueba para asegurarse que se pueda realizar la petición partiendo desde la barra de búsqueda	
<b>Condiciones de ejecución:</b> El usuario introduce el dato deseado	
<b>Pasos de ejecución:</b> El usuario debe introducir en el campo de la barra de búsqueda el parámetro deseado para posteriormente realizar la búsqueda	
<b>Resultados esperados:</b> Una vez introducida la información que se desea buscar, de coincidir con las hojas existentes en el árbol, se realiza la búsqueda de la información	

**Tabla 62. Prueba de aceptación#19**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU14_P2	<b>Historia de Usuario:</b> 14
<b>Nombre:</b> Realizar petición RSS partiendo del Barra de Búsqueda	
<b>Descripción:</b> Prueba para asegurarse que se pueda realizar la petición partiendo desde la barra de búsqueda	
<b>Condiciones de ejecución:</b> El usuario introduce el dato deseado	
<b>Pasos de ejecución:</b> El usuario debe introducir en el campo de la barra de búsqueda el parámetro deseado para posteriormente realizar la búsqueda	

**Resultados esperados:** Una vez introducida la información que se desea buscar, si no coincide con las hojas del árbol muestra el mensaje “Los parametros no son correctos”

**Tabla 63. Prueba de aceptación#20**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU15_P1	<b>Historia de Usuario:</b> 15
<b>Nombre:</b> Cargar Resultados RSS	
<b>Descripción:</b> Se realiza la búsqueda de la información deseada	
<b>Condiciones de ejecución:</b> Cargar el listado que muestra los resultados obtenidos	
<b>Pasos de ejecución:</b> Una vez introducido los datos seleccionados se realizan la llama a la API, para la búsqueda de la información	
<b>Resultados esperados:</b> De ser satisfactoria la búsqueda se muestra el listado con los resultados obtenidos	

**Tabla 64. Prueba de aceptación#21**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU15_P2	<b>Historia de Usuario:</b> 15
<b>Nombre:</b> Cargar Resultados RSS	
<b>Descripción:</b> Se realiza la búsqueda de la información deseada	
<b>Condiciones de ejecución:</b> Cargar el listado que muestra los resultados obtenidos	
<b>Pasos de ejecución:</b> Una vez introducido los datos seleccionados se realizan la llama a la API, para la búsqueda de la información	
<b>Resultados esperados:</b> Muestra el mensaje “No se pudo obtener resultados”	

**Tabla 65. Prueba de aceptación#22**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU16_P1	<b>Historia de Usuario:</b> 16
<b>Nombre:</b> Filtrar Resultados RSS	
<b>Descripción:</b> Se seleccionan los datos por los que se desea filtrar la información	
<b>Condiciones de ejecución:</b> Seleccionar los atributos de filtrado	
<b>Pasos de ejecución:</b> El usuario selecciona los atributos de filtrado deseados, como la cantidad de resultados a mostrar en pantalla y el rango de años de los mismos	
<b>Resultados esperados:</b> Se cargan los resultados cumpliendo con los atributos de filtrado seleccionados	

### 3.4.4 Iteración IV

**Tabla 66. Prueba de aceptación#23**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU17_P1	<b>Historia de Usuario:</b> 17
<b>Nombre:</b> Realizar petición Noticias partiendo del TreeView	
<b>Descripción:</b> Prueba para asegurarse que se puedan escoger los datos a buscar desde el árbol	
<b>Condiciones de ejecución:</b> El usuario debe seleccionar la hoja que represente la información que se desea	
<b>Pasos de ejecución:</b> El usuario debe desplegar las ramas del árbol que le interesan hasta llegar a las hojas y seleccionar la hoja con la información que se desea buscar	
<b>Resultados esperados:</b> Una vez seleccionada la información deseada se envía al backend y se procede a realizar la búsqueda de los resultados	

**Tabla 67. Prueba de aceptación#24**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU18_P1	<b>Historia de Usuario:</b> 18
<b>Nombre:</b> Realizar petición Noticias partiendo del Barra de Búsqueda	
<b>Descripción:</b> Prueba para asegurarse que se pueda realizar la petición partiendo desde la barra de búsqueda	
<b>Condiciones de ejecución:</b> El usuario introduce el dato deseado	
<b>Pasos de ejecución:</b> El usuario debe introducir en el campo de la barra de búsqueda el parámetro deseado para posteriormente realizar la búsqueda	
<b>Resultados esperados:</b> Una vez introducida la información que se desea buscar, de coincidir con las hojas existentes en el árbol, se realiza la búsqueda de la información	

**Tabla 68. Prueba de aceptación#25**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU18_P2	<b>Historia de Usuario:</b> 18
<b>Nombre:</b> Realizar petición Noticias partiendo del Barra de Búsqueda	
<b>Descripción:</b> Prueba para asegurarse que se pueda realizar la petición partiendo desde la barra de búsqueda	
<b>Condiciones de ejecución:</b> El usuario introduce el dato deseado	
<b>Pasos de ejecución:</b> El usuario debe introducir en el campo de la barra de búsqueda el parámetro deseado para posteriormente realizar la búsqueda	
<b>Resultados esperados:</b> Una vez introducida la información que se desea buscar, si no coincide con las hojas del árbol muestra el mensaje "Los parametros no son correctos"	

**Tabla 69. Prueba de aceptación#26**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU19_P1	<b>Historia de Usuario:</b> 19
<b>Nombre:</b> Cargar Noticias	
<b>Descripción:</b> Se realiza las busqueda de la información deseada	
<b>Condiciones de ejecución:</b> Cargar el listado que muestra los resultados obtenidos	
<b>Pasos de ejecución:</b> Una vez introducido los datos seleccionados se realizan la llama a la API, para la búsqueda de la información	
<b>Resultados esperados:</b> De ser satisfactoria la búsqueda se muestra el listado con los resultados obtenidos	

**Tabla 70. Prueba de aceptación#27**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU19_P2	<b>Historia de Usuario:</b> 19
<b>Nombre:</b> Cargar Resultados RSS	
<b>Descripción:</b> Se realiza la búsqueda de la información deseada	
<b>Condiciones de ejecución:</b> Cargar el listado que muestra los resultados obtenidos	
<b>Pasos de ejecución:</b> Una vez introducido los datos seleccionados se realizan la llama a la API, para la búsqueda de la información	
<b>Resultados esperados:</b> Muestra el mensaje "No se pudo obtener resultados"	

**Tabla 71. Prueba de aceptación#28**

<b>Caso de prueba de aceptación</b>	
<b>Código:</b> HU20_P1	<b>Historia de Usuario:</b> 20
<b>Nombre:</b> Ver Detalles	
<b>Descripción:</b> Permite al usuario obtener más información de los resultados obtenidos	
<b>Condiciones de ejecución:</b> Haber realizado una búsqueda anteriormente	
<b>Pasos de ejecución:</b> Se selecciona el resultado de búsqueda obtenido del cual se deseen ver más detalles y se selecciona el botón ver	

**Resultados esperados:** Muestra en pantalla una vista con los detalles de los resultados

## **Conclusiones Parciales**

A lo largo de este capítulo se abordaron los temas principales de la implementación del sistema y las pruebas de aceptación realizadas. En primera instancia se describieron un total de 20 tareas de ingeniería, planteando las acciones a seguir por el equipo de desarrollo para conformar el sistema para la búsqueda y recopilación de información. Posteriormente se realizaron las pruebas aceptación. De estas se pudo obtener un total que se realizaron 28 pruebas de aceptación, siendo 22 satisfactorias, demostrando de este modo el cumplimiento de los requisitos fundamentales propuestos.

# Conclusiones

Con este trabajo se desarrolló un sistema de búsqueda y recopilación de información científica para la implementación de un Plataforma Web para la Vigilancia Tecnológica, llevado a cabo por el MINCOM y la UCI. El sistema propuesto utiliza diversas tecnologías como las API, los canales RSS, y el uso del Webscraping, para la obtención de información científica y académica, relacionada con las ciencias de la computación. Adicionalmente el sistema incluye:

- Una vista de árbol que contiene las distintas categorías y temas derivados de las ciencias de la computación, constituyendo de este modo una guía para los usuarios, ayudándolos a llegar a la información que desean.
- Los resultados obtenidos cuentan con un nivel de relevancia o rating, representados mediante estrellas. De este modo permite a los usuarios conocer que tan bueno es considerado dicha información o no, puesto que este rating se determina por el número de citas presentadas por el documento y según el promedio del rating introducido por todos los usuarios.
- Este sistema permite de manera indirecta una comunicación y retroalimentación entre los distintos usuarios, no solo utilizando el índice asociado a las estrellas, sino también a un sistema que permite a los usuarios añadirle a cada resultado obtenido un comentario. De este modo se le permite a los usuarios conocer la opinión del resto sobre un resultado de búsqueda en específico.

# Bibliografía

1. (2012) “La vigilancia tecnológica y la inteligencia competitiva en los estándares de la gestión de calidad I+D+I” from [https://adingores.sserver.es/congresos/web/uploads/cio/cio2012/SP\\_04\\_Gestion\\_Innovacion\\_Tecnologica\\_y\\_Organizativa/1162-1168.pdf](https://adingores.sserver.es/congresos/web/uploads/cio/cio2012/SP_04_Gestion_Innovacion_Tecnologica_y_Organizativa/1162-1168.pdf)
2. “¿Qué es la Vigilancia Tecnológica? Definición y desafíos” from <https://datascientest.com>
3. (2020) “Guía de Vigilancia e Inteligencia Tecnológica — OVTT” from <https://www.ovtt.org/guias/guia-de-inteligencia-tecnologica>
4. “La VT en Cuba - La Vigilancia Tecnológica Fundamentos teóricos” from <https://1library.co/article/vt-cuba-vigilancia-tecnol%C3%B3gica-fundamentos-te%C3%B3ricos.y4qj710y>
5. San Juan, Yelena Islen; Romero Rodríguez, Felix Ivan." Modelos y herramientas para la vigilancia tecnológica in Ciencias de la Información, vol. 47, no. 2" Instituto de Información Científica y Tecnológica La Habana, 2016
6. Jiménez Jiménez, Daniel. Martínez Costa, Micaela. Marín Baquero, Diego. “La vigilancia tecnológica y la cooperación para el desarrollo de la innovación: una aproximación empírica” Universidad de Murcia, España, 2017
7. “SINFORCIT | Vigilancia Tecnológica” from <http://www.sinforcit.sld.cu/site/vigtec.html>

8. "Vigilancia Tecnológica - Ruta N" from <https://www.rutanmedellin.org/es/recursos/abc-de-la-innovacion/item/vigilancia-tecnologica>
9. . Fernando Martínez Rivero, Elsa Rosana Maynegra Díaz "Evaluación de plataformas web para su implementación en el sistema de vigilancia tecnológica de la Consultoría Biomundi", Rev. cuba. inf. cienc. salud vol.25 no.1 La Habana mar, 2014
10. (2021) "XerkaCv" from <https://xerkariak.com/es/xerka-cv-cas/>
11. "Vixía" Vigilancia Tecnológica Y Del Entorno +gestión De La - ID:5f062dd734285" from <https://xdoc.mx/documents/vixia-vigilancia-tecnologica-y-del-entorno-gestion-de-la-5f062dd734285>
12. (2021)Plataforma VIGIALE®" from <https://www.vigiale.com/>
13. "Bienvenido a Vicubo Cloud. Software de Inteligencia Competitiva y Vigilancia Tecnológica" from <https://www.vicubocloud.es/es/node>
14. "SOFTVT - Herramienta de Vigilancia Tencnológica" from <https://www.softvt.com/>
15. "Miniera | LinkedIn" from <https://es.linkedin.com/company/miniera>
16. "Hontza | Hontza es una Plataforma de Software Libre para la Inteligencia Competitiva y Estratégica" from <http://www.hontza.es/>
17. Omar Correa Madrigal "Vigilancia Científico-Técnica en TIC con uso intensivo de tecnologías computacionales", Universidad de las Ciencias Informáticas

18. María Isabel Ramírez, David Escobar Rúa, Bibiana Arango Alzate (2012) "Vigilancia Tecnológica e Inteligencia Competitiva"
19. "CDE - Vigilancia Tecnológica e Inteligencia Competitiva" from [https://www.cde.es/es/inteligencia\\_competitiva/](https://www.cde.es/es/inteligencia_competitiva/)
20. Raudel Giraldez Reyes, Maidelyn Díaz (2018) "Observatorio de artículos científicos en Acceso Abierto" from <http://www.congreso-info.cu/index.php/info/info2018/paper/view/868/542>
21. Manuel Blázquez Ochando (2012) "Técnicas avanzadas de recuperación de información"
22. R. Baeza Yates, B. Ribeiro Neto (1999) "Modern Information Retrieval"
23. G. Salton (1989) "Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer."
24. R. Sanchez Jimenez (2011) "Técnicas avanzadas de recuperación de información"
25. S. Levithan, J. Goyvaert (2009) "Regular Expressions Cookbook OReilly" from <http://www.bookf.net/p/3844>)
26. G. Salton (1983) "Introduction to Modern Information Retrieval"
27. F.W Lancaster, A.J Warner (1993) "Information Retrieval Today. Information Resources"
28. Grela, L. (2020). "Vocabulario controlado vs Vocabulario libre." from [https://blog.uchceu.es/biblioteca/vocabulario-controlado-vs-vocabulario-libre/?\\_adin=02021864894](https://blog.uchceu.es/biblioteca/vocabulario-controlado-vs-vocabulario-libre/?_adin=02021864894)
29. Andrea Núñez-Torrón(2021)"¿Qué es una API y para qué sirve? Todo lo que necesitas saber" from <https://www.businessinsider.es/api-sirve-todo-necesitas-saber-861403>.
30. Hat, R. (2020). "¿Qué es una API de REST?". from <https://www.redhat.com/es/topics/api/what-is-a-rest-api>.
31. Hat, R. (2020). "Diferencias entre REST y SOAP." from <https://www.redhat.com/es/topics/integration/whats-the-difference-between-soap-rest>

32. "¿Qué es un feed RSS?" from <https://www.microsiervos.com/archivo/general/que-es-un-feed-rss.html>
33. Danny Murillo, Dalys Saavedra " Web Scraping de los Perfiles y Publicaciones de una Afiliación en Google Scholar utilizando Aplicaciones Web e implementando un Algoritmo en R", VIPE, Universidad Tecnológica de Panamá ,2017
34. (2020)"Web scraping con Python. Guía de inicio de Beautiful Soup" from <https://j2logo.com/python/web-scraping-con-python- guia-inicio-beautifulsoup/>
35. IONOS (2020). "Web scraping con Python: introducción y tutorial." from <https://www.ionos.es/digitalguide/paginas-web/desarrollo-web/web-scraping-con-python.>
36. ETI "Plataforma Vigintel" from <https://www.eti.cu>
37. (2022) "DFL Observatorio: una herramienta ideal para gestionar la información" from <https://www.dofleini.com/dfl-observatorio-una-herramienta-ideal-para-gestionar-la-informacion/>
38. Maida EG, Pacienza J (2015) "Metodologías de desarrollo de software [en línea]. Tesis de Licenciatura en Sistemas y Computación. Facultad de Química e Ingeniería "Fray Rogelio Bacon". Universidad Católica Argentina" from <https://repositorio.uca.edu.ar/bitstream/123456789/522/1/metodologias-desarrollo-software.pdf>
39. Milena Meléndez, María Gaitan, Neldin Pérez (2016) "Metodología Ágil de Desarrollo de Software Programación Extrema" from <https://repositorio.unan.edu.ni/1365/1/62161.pdf>
40. "Metodología XP o Programación Extrema" from <https://www.sinnaps.com/blog-gestion-proyectos/metodologia-xp>
41. "Python: qué es y por qué deberías aprender a utilizarlo" from <https://www.becas-santander.com/es/blog/python-que-es.html>
42. (2013) "Qué es Python" from <https://desarrolloweb.com/articulos/1325.php>
43. "Funcionalidades - PyCharm" from <https://www.jetbrains.com/es-es/pycharm/features/>

44. "Compatibilidad de PyCharm" from <https://www.componentsource.com/es/product/pycharm/compatibilities>
45. "¿cuáles Son Las Desventajas De Visual Studio?" from <https://mejorsoftware.info/app33/40656/cuales-son-las-desventajas-de-visual-studio>
46. "¿cuáles Son Las Características De Visual Studio?" from <https://mejorsoftware.info/app33/7264/cuales-son-las-caracteristicas-de-visual-studio>
47. "Microsoft Visual Studio, Concepto y ¿Qué es y para qué sirve Microsoft Visual Studio?" from <https://www.espaciohonduras.net/microsoft-visual-studio-concepto-y-que-es-y-para-que-sirve-microsoft-visual-studio>
48. "The web framework for perfectionists with deadlines | Django" from <https://www.djangoproject.com/>
49. C. Acuña (2017) "Arquitecturas de Software" from <https://www.google.com/search?client=firefox-b-d&q=C.+Acu%C3%B1a,+%22Arquitecturas+de+Software%22,Pontificia+Universidad+Javeriana,2017.%5BOnline%5D.Available:http://pegasus.javeriana.edu.co/~mad/Arquitecturas%2520de%2520SW.pdf&spell=1&sa=X&ved=2ahUKEwi-4iGhtn6AhX2STABHfmQAhcQBSgAeqQIBhAB&biw=1600&bih=775>
50. "Estilos Arquitectónicos en la Ingeniería de Software" from <https://www.studocu.com/co/document/universidad-pedagogica-y-tecnologica-de-colombia/ingenieria-de-software/estilos-arquitectonicos-en-la-ingenieria-de-software/5107767>
51. Yván Ecarri (2016) "Estilos y Patrones de Arquitectura de Software" from <https://blog.osoft.es/index.php/2016/03/24/estilos-y-patrones-de-arquitecturas-de-software/>
52. "Arquitectura Cliente-Servidor" from <https://reactiveprogramming.io/blog/es/estilos-arquitectonicos/cliente-servidor>
53. "Arquitectura cliente-servidor: definición, tipos, ejemplos, ventajas y desventajas" from <https://aprendiendoaprogramar.es/blog/arquitectura-cliente-servidor-definicion-tipos-ejemplos-ventajas-y-desventajas/>