



**INSTITUTO TECNOLÓGICO SUPERIOR  
DE MISANTLA**

---

**SUBSISTEMA DE SEGURIDAD PARA EL  
CONTROL DE ACCESO Y TRANSFERENCIA  
DE DATOS DEL SITM.**

**TESIS**

PARA OBTENER EL GRADO DE MAESTRÍA EN  
SISTEMAS COMPUTACIONALES

**P R E S E N T A**

FERNANDO ALBERTO HERNANDEZ GUEVARA

**DIRECTOR**

MSC. GALDINO MARTINEZ FLORES

**CO-DIRECTOR**

DR. JORGE MARIO FIGUEROA GARCÍA

MISANTLA, VERACRUZ

MARZO, 2020



# Autorización de impresión

 INSTITUTO TECNOLÓGICO SUPERIOR DE MISANTLA  
DIVISIÓN DE ESTUDIOS PROFESIONALES  
AUTORIZACIÓN DE IMPRESIÓN DE TRABAJO DE TITULACIÓN MAESTRÍA

FECHA: 19 de Marzo de 2020

ASUNTO: AUTORIZACIÓN DE IMPRESIÓN DE TESIS.

**A QUIEN CORRESPONDA:**

Por medio de la presente se hace constar que el (la) C:

FERNANDO ALBERTO HERNÁNDEZ GUEVARA

estudiante de la maestría en SISTEMAS COMPUTACIONALES con No. de Control 172T0858 ha cumplido satisfactoriamente con lo estipulado por el Lineamiento de Posgrado para la obtención del grado de Maestría mediante Tesis.

Por tal motivo se Autoriza la impresión del Tema titulado:

**SUBSISTEMA DE SEGURIDAD PARA EL CONTROL DE ACCESO Y TRANSFERENCIA DE DATOS DEL SITM**

Dándose un plazo no mayor de un mes de la expedición de la presente a la solicitud del examen para la obtención del grado de maestría.

ATENTAMENTE

  
MSC. Galdino Martínez Flores  
Presidente

 DIVISIÓN DE ESTUDIOS PROFESIONALES

  
MSC. Eddy Sánchez de la Cruz  
Secretario

  
Dr. Jorge Mario Figueroa García  
Vocal

Archivo.

VER. 01/03/09 F-SA-39

---

# Dedicatoria

*Dedicado  
con todo mi amor y cariño  
a mi amada esposa Keila E.O.D.  
por su sacrificio y esfuerzo, por creer en mi capacidad,  
siempre has estado brindándome tu  
comprensión, apoyo, cariño y amor.*

*A mi amado hijo Fernando H.O.  
por ser mi fuente de motivación e inspiración para  
poder superarme cada día más y así poder luchar  
para que la vida nos depare un futuro mejor.*

*A mi amado padre y hermana  
quienes con su amor no me dejaban caer.*

*Gracias a todos.*

---

# Agradecimientos

Dios, tu amor y tu bondad no tienen fin, me permites sonreír ante todos mis logros que son resultado de tu ayuda, y cuando caigo y me pones a prueba, aprendo de mis errores y me doy cuenta que los pones en frente mío para que mejore como ser humano, y crezca de diversas maneras.

A mi familia, keila, fer, hermana y papá, porque con sus oraciones, consejos y palabras de aliento hicieron de mí una mejor persona y de una u otra forma me acompañan en todos mis sueños y metas.

A mis amigos Marcela, Ana, Ángel, Gamaliel, Luis Enrique, Iran, Carlos, Isidro, Compañeros y Amigos que hacen la comunidad del ITSM, por su cariño y apoyo incondicional, durante todo este proceso, por estar conmigo en todo momento gracias. Agradezco a mis profesores, Doctores y Maestros de la Maestría en Sistemas Computacionales de ITSM, por haber compartido sus conocimientos a lo largo de mi preparación en esta profesión.

Finalmente quiero expresar mi más grande y sincero agradecimiento al Dr. Jorge Mario, principal colaborador durante todo este proceso, quien con su dirección, conocimiento, enseñanza y colaboración permitió el desarrollo de este trabajo.

Es difícil mencionar a todos(as) ustedes, que estuvieron apoyándome en cada parte de este proceso, que ha sido el cursar la maestría.

¡Muchas gracias a todos!

---

# Resumen

La presente investigación muestra la implementación de una Prueba de Turing completamente automática y pública para diferenciar ordenadores de humanos con Algoritmos Genéticos, con la finalidad de facilitar el acceso a los usuarios y a su vez garantizar que sea un humano el que trata de ingresar al Sistema Integral del Tecnológico de Misantla. El documento pretende concientizar a los programadores de sistemas de información que desarrollan en un entorno web, que no basta con terminar el mismo, sino que deben de estar conscientes de que en los sistemas web existen más riesgos que en los sistemas desktop, por la naturaleza pública y multiplataforma que estos pueden ser, estos podrían intentar ser vulnerados por ciber atacantes. También a su vez que las técnicas o metodologías de seguridad que se desarrollan no sea una molestia o algo tedioso para el usuario final y que consuma muchos recursos de los equipos de los clientes y de los servidores donde se encuentran alojados.

---

# Abstract

the present investigation sample the implementation of a Completely Automated Public Turing test, to differentiate computers from humans with Genetic Algorithms and thus facilitate access for users to Misantla Technology Integral System (“Sistema Integral del Tecnológico de Misantla”). The article Its aims to raise awareness among information system programmers who develop in a web environment, about the importance of establishing security tools for access to them, because the risks are greater compared to desktop systems, derived from its public and cross-platform nature and in turn the possible vulnerability to cyber attackers. Also, that the security techniques or methodologies that are developed are not a nuisance or something tedious for the end user and that it consumes too many resources from its computers and from the servers where they are hosted.

---

# Índice general

Autorización de impresión	I
Dedicatoria	II
Agradecimientos	III
Resumen	IV
Abstract	V
Índice de figuras	IX
<b>1. Generalidades</b>	<b>1</b>
1.1. Introducción . . . . .	2
1.2. Motivación . . . . .	3
1.3. Justificación . . . . .	3
1.4. Planteamiento del Problema . . . . .	3
1.5. Propuesta de Solución . . . . .	4
1.6. Hipótesis . . . . .	4
1.7. Objetivos . . . . .	4
1.7.1. Objetivo General . . . . .	4
1.7.2. Objetivos Especificos . . . . .	5
1.8. Alcances y Limitaciones . . . . .	5
1.8.1. Alcances . . . . .	5
1.8.2. Limitaciones . . . . .	6
	VI



---

<b>2. Marco Teórico</b>	<b>8</b>
2.1. BackEnd . . . . .	11
2.1.1. Php . . . . .	11
2.1.2. MySQL . . . . .	12
2.1.3. XML . . . . .	13
2.1.4. Base de Datos . . . . .	13
2.2. FrontEnd . . . . .	14
2.2.1. Dom . . . . .	14
2.2.2. Html . . . . .	15
2.2.3. Diseño Web Adaptativo . . . . .	15
2.2.4. ECMAScript . . . . .	16
2.2.5. Html5 . . . . .	16
2.2.6. JavaScript . . . . .	17
2.3. Entornos de desarrollo . . . . .	17
2.3.1. Linux . . . . .	17
2.3.2. Apache . . . . .	19
2.3.3. Sublime Text . . . . .	19
2.3.4. Xampp . . . . .	20
2.4. Intercambio de datos . . . . .	21
2.4.1. Ajax . . . . .	21
2.4.2. JSON . . . . .	21
2.5. Seguridad informática . . . . .	22
2.5.1. Hacker . . . . .	22
2.5.2. Spyware . . . . .	22
2.5.3. Troyano . . . . .	23
2.5.4. Phishing . . . . .	23
2.5.5. Bontnets . . . . .	24
2.5.6. Trashing . . . . .	24
2.6. Programación Orientada a Objetos (POO) . . . . .	24
2.6.1. Características conceptuales de la POO . . . . .	25
2.7. Patrón de arquitectura de Software. . . . .	26
2.7.1. Modelo Vista Control (MVC) . . . . .	26

---

2.8. Modelo de procesos del software . . . . .	27
2.8.1. Modelo en espiral: . . . . .	27
2.9. Técnicas de aprendizaje automático . . . . .	28
2.9.1. Algoritmos Genéticos (AGs) . . . . .	28
<b>3. Estado del Arte</b>	<b>30</b>
<b>4. Metodología</b>	<b>37</b>
4.1. Algoritmo Genético . . . . .	39
4.1.1. Codificación de los genes . . . . .	40
4.1.2. Población inicial . . . . .	40
4.1.3. Selección . . . . .	41
4.1.4. Cruza . . . . .	41
4.1.5. Mutación . . . . .	42
4.1.6. Resumen de parametros del AG . . . . .	42
<b>5. Resultados y Análisis</b>	<b>43</b>
5.1. Descripción de la Interfaz . . . . .	44
5.2. Descripción del código . . . . .	48
<b>6. Conclusiones y Trabajos futuros</b>	<b>54</b>
6.1. Conclusiones . . . . .	55
6.2. Trabajos futuros . . . . .	56

---

# Índice de figuras

2.1. Modelo en espiral . . . . .	27
2.2. Pseudo código del Algoritmo Genético Simple . . . . .	29
4.1. Diseño gráfico de la metodología. . . . .	38
5.1. Integración AG. . . . .	45
5.2. Ingreso de la palabra en AG. . . . .	45
5.3. Corrida del AG. . . . .	46
5.4. Se habilita en botón ingresar. . . . .	46
5.5. Logeo satisfactorio. . . . .	47
5.6. Error AG. . . . .	47

---

# Índice de cuadros

4.1. Resumen de los parámetros del algoritmo evolutivo. . . . .	42
---	----

---

# Capítulo 1

## Generalidades

## 1.1. Introducción

Los sistemas de información desarrollados en un entorno web representan una tecnología en crecimiento, entre las cuales se pueden destacar su fácil acceso desde cualquier parte del mundo, siempre y cuando el usuario final cuente con un navegador de internet, gracias a esto, no se limita solo a un equipo de cómputo, como es bien sabido los dispositivos móviles, ya sean smartphones y/o tabletas tienen dicha conexión. Lo primero no solo representa ventajas, sino también inconvenientes, al estar expuesto de forma pública, ya que existen personas mal intencionadas que buscan vulnerar dichos sistemas, para robar información. No solo humanos son los que intentan penetrar en los sistemas web, también los algoritmos o programas maliciosos buscan saturar los servidores donde se encuentran alojados dichos sistemas, saturándolos con peticiones de información, tanto en formularios de inicio de sesión, solicitud de información, contacto, entre otros. Existen métodos para garantizar la seguridad e integridad de los sistemas de información web.

Actualmente las pruebas de Turing o captcha implementadas representan una herramienta útil, pero en ocasiones suelen ser molestas para los usuarios, debido a que tienden a ser procesos extensos, repetitivos y tediosos, lo que resulta difícil o hasta incómodo, al contrario de lo que los programadores pueden pensar que es sencillo e idóneo para resolver un problema al usuario final, no lo es.

El Sistema Integral del Tecnológico de Misantla (SITM) está desarrollado para un entorno web, es por ello que, puede ser objeto de los ataques antes mencionados. El hecho de que contenga módulos que, por su configuración, requieran ser visualizados/consultados desde fuera de la institución, en cierto grado de seguridad lo hace vulnerable a posibles ataques. Por mencionar algunos de los módulos que pueden ser atacados se encuentra, el acceso al sistema, la transferencia y el almacenamiento de la información.

El objetivo, es una propuesta de solución para mitigar la problemática que pudieran ocasionar los ciber atacantes al SITM, en el aspecto de acceder a los módulos del sistema y/o iniciar sesión, implementando una herramienta que garantice que quién intenta acceder sea un humano o persona real y no un algoritmo de spam que sature el servidor con solicitudes de sesión.

## **1.2. Motivación**

Considerar que será de gran aporte al sistema que actualmente se desarrolla en el departamento de sistematización y redes del Tecnológico de Misantla, SITM(Sistema Integral del Tecnológico de Misantla).

## **1.3. Justificación**

Es necesario diseñar, implementar e implantar el subsistema de seguridad para el control de acceso y transferencia de datos del Sistema Integral del Tecnológico de Misantla, para fortalecer la seguridad del sistema de información con la ayuda de un protocolo SSL y prueba de Turing automática y pública para diferenciar ordenadores de humanos, garantizara el acceso seguro y transferencia de la información del sistema integral del tecnológico de Misantla. Lo anterior debido a que la seguridad de los sistemas de información web es primordial para su correcto funcionamiento y que no existan perdida, duplicidad, espionaje o intrusión en la información manejados durante el proceso de captura de los datos y a su vez que la información existente en la base de datos se mantenga integra.

## **1.4. Planteamiento del Problema**

En la actualidad el uso de los servicios de información web ofrece una amplia gama de oportunidades comerciales y no comerciales, tanto en el sector privado como el público, debido a dicho auge de estas tecnologías, se abren las puertas a los ciber ataques, la web no es una excepción a este patrón, no están exentos de ataques o de personas malintencionadas que solo buscan su propio provecho. Hoy, millones de ciber atacantes están avanzando en una amplia variedad de estafas que incluyen los dominios de propagación. Siendo el SITM (Sistema Integral del Tecnológico de Misantla) un sistema de información desarrollado para un entorno web, es por ello que, puede ser objeto de los ataques antes mencionados. El hecho de que contenga módulos que por su configuración requieran ser visualizados/consultados desde fuera de la institución, en cierto grado de seguridad lo hace vulnerable a posibles ataques. Por mencionar algunos de los módulos que pueden ser atacados se encuentra, el acceso

al sistema, la transferencia y el almacenamiento de la información. Es por lo anterior que el presente trabajo busca una propuesta de solución para mitigar la problemática que los ciber ataques que pudieran ocasionar problemas al SITM, en los aspectos de acceder a los módulos del sistemas, transferir la información de forma íntegra de tal forma que los datos no se pierdan o sean intervenidos en la comunicación que se abre entre el cliente y el servidor, y por ultimo proveer de seguridad al almacenamiento de datos, sin que estos puedan ser penetrados o corrompidos por usuarios no autorizados, ajenos al uso del sistema de información.

## **1.5. Propuesta de Solución**

Los certificados SSL son una tecnología probada y con gran uso en los sistemas de información web actuales ya que ayudan a mantener una comunicación cifrada y encriptada entre el cliente y el servidor donde se encuentra alojado evitando ataques de un experto en informática donde pudiera colocarse como intermediario y así capturar la información que se está enviando entre los usuarios del SITM y el Servidor donde se encuentra el SITM. A su vez una prueba de Turing automática y pública para diferenciar ordenadores de humanos, garantizara el acceso seguro y transferencia de la información del sistema integral del tecnológico de Misantla.

## **1.6. Hipótesis**

La aplicación de un protocolo SSL y prueba de Turing automática y pública para diferenciar ordenadores de humanos, garantizará el acceso seguro y transferencia de la información del sistema integral del tecnológico de Misantla.

## **1.7. Objetivos**

### **1.7.1. Objetivo General**

Implementar el subsistema de seguridad para el control de acceso y transferencia de datos de manera segura e íntegra al Sistema Integral del Tecnológico de Misantla.



### 1.7.2. Objetivos Especificos

- Investigar los sistemas de seguridad implementadas en los diferentes servidores web mediante el análisis de documentos como son, artículos, libros y revistas, todos de carácter científico, para garantizar que la información viaje de forma segura e integra entre el cliente y el servidor.
- Diseñar la metodología de seguridad para Sistema Integral del Tecnológico de Misantla(SITM) mediante el modelo en espiral por su facilidad de cambio al momento de desarrollar propuestas de solución.
- Desarrollar el módulo de contro de acceso usando algoritmos genéticos para realizara la prueba de Turing para garantizar que sea un humano quien es el que intenta acceder al sistema de información y no un algoritmo mal intencionado, que pudiera generar spam saturando el servicio, y con esto ocasionar un problema mayor.
- Integrar el módulo de seguridad al SITM mediante la integración del algoritmo genético y el certificado SSL para fortificar la seguridad del SITM.
- Validación de resultados mediante la generación de pruebas de ingreso al SITM para comprobar que la propuesta de solución implantada cumple con su finalidad de reforzar el sistema de información.

## 1.8. Alcances y Limitaciones

### 1.8.1. Alcances

Para el desarrollo del módulo de seguridad se planean alcanzar las siguientes metas:

- Escaneo de vulnerabilidades del Sistema Integral del Tecnológico de Misantla. Esto para identificar los puntos vulnerables del sistema mediante un análisis exhaustivo con herramientas de prueba de penetración, mismos que proveerán resultados que amplíen la situación actual del dicho y con ello ayudar a la toma de decisiones con respecto a las soluciones a implementar.

- Implementación de un certificado SSL. con esto se crea un canal seguro, donde la información viajara de forma cifrada, ya que, al cambiar el contexto de la información, será de difícil compresión para el atacante que intente intervenir la comunicación entre el cliente y el servidor.
- Desarrollo de un y prueba de Turing automática y pública para diferenciar ordenadores de humanos. mediante la integración de los algoritmos genéticos para proveer de un metodo mas seguro a dicha prueba.
- Garantizará el acceso seguro y transferencia de la información del sistema integral del tecnológico de Misantla. Lo anterior con el certificado de navegación segura (SSL) y el desarrollo de una prueba de Turing con Algoritmos Genéticos se fortalecerá la seguridad del sistema, que la información viaje de forma íntegra y controlada hacia el SITM.

### 1.8.2. Limitaciones

- Los recursos de hardware con los que cuenta el servidor actual no cumple con las caracteristica necesarias para la implantación de la propuesta de solución para ello, el equipo debe de contar con un Procesador Xeon, este tipo de procesador soporta altas temperatura y un trabajo 24/7 debido a su composición, Memoria RAM 16 GB como mínimo para soportar las tareas y actividades que se ejecuten en el servidor, 2 Discos Duros 1TB con soportar de sistema RAID para una escritura redundante y tener en todo momento un respaldo, enfriamiento liquido con esto la temperatura del equipo se mantendrá operable sin calentamiento excesivo, un sistema operativo basado en Linux como es Ubuntu Server, con dicho sistema se ahorrar en el gasto innecesario de licencias y disminuir la probabilidad de virus, por ultimo y no menos importante, es que debe de contar con conexión de Internet con una dirección IP publica para que el sistema puedan acceder desde cualquier parte del mundo, que de igual manera tenga acceso a la red universal..
- Los costos elevados de los Certificados SSL representan sin lugar a duda representan una limitante, ya que los recursos con los que se cuentan son limitados por no decir inexistentes, a mayor seguridad mayor costo.

- 
- Escasa información existente de trabajos relacionados con este proyecto, los trabajos relacionados a esta investigación en el campo de la seguridad web con la integración de un algoritmo genético y una prueba de Turing es insuficiente.
  - Los lineamientos impuestos por instituciones superiores, como son el tecNM, quien a forzado a las instituciones a ocupar una plantilla y un certificado, que sin lugar a duda se desconoce el nivel de seguridad en cuanto al cifrado que puede proponer para los fines de dicho sistema.
  - La curvatura de aprendizaje es alta, ya que el tiempo con el que se cuenta es poco ya que el aprender de múltiples lenguajes de programación no es difícil pero, se debe de tener una comprensión alta para la generación de la solución mediante la integración de los lenguajes debe ser meticuloso.

---

## Capítulo 2

### Marco Teórico

En la actualidad los sistemas de información desarrollados en un entorno web representa una tecnología que está en crecimiento ya que sus ventajas representan mayores beneficios para las empresas, gracias a que al estar en un entorno orientado a la web, se puede acceder desde cualquiera lugar, ya sea localmente dentro de la misma empresa o desde fuera o parte del mundo, siempre que se tenga acceso a internet, un navegador web actualizado que soporte las tecnologías en que fue programado y las políticas de la empresa así lo permitan, a diferencia de las aplicaciones que están hechas para un entorno de escritorio.

Aunque esto trae múltiples inconvenientes ya que no está exento a un gran número de vulnerabilidades, debido a que, al estar expuesta de cara al público, puede sufrir múltiples ataques que pueden ocasionar desde robo de información, pérdida de datos, inconsistencias en los registros hasta dejar imposibilitado los sistemas de información.

Es por lo anterior que las empresas invierten gran parte de su presupuesto para mitigar las posibles vulnerabilidades que puedan tener los sistemas que desarrollan, contratando a expertos en la materia de pruebas de penetración a sistemas para que estas generen un reporte a detalle con todos los puntos más vulnerables en sus aplicaciones.

Actualmente existen herramientas que permiten detectar de forma automática los problemas de desarrollo y vulnerabilidad en los sistemas desarrollados en un entorno web, un ejemplo de ello es Acunetix Web Vulnerability Scanner que permite realizar una auditoría de la seguridad web e identificar si el sitio es hackeable por los piratas informáticos.(«Acunetix Vulnerability Scanner», s. f.)

Acunetix son los pioneros en las pruebas de seguridad de aplicaciones web automatizadas con tecnologías innovadoras que incluyen:

- Tecnología DeepScan : para el rastreo de aplicaciones de página única (SPA) pesadas en el lado del cliente de .
- La prueba más avanzada de SQL Injection y Cross-site Scripting incluye detección avanzada de XSS basado en DOM.
- Tecnología AcuSensor: combinas técnicas de escaneo de caja negra con retroalimentación de sus sensores colocados dentro del código fuente.

Otra herramienta que ayuda a identificar puertas traseras que pudieran ocasionar problemas a los sistemas de información debido a al exceso de puertos abiertos en el servidor que se encuentra implantado es NMAP.

NMAP es un programa de código abierto, actualmente multiplataforma, que ayuda a realizar un rastreo de puertos, se usa para evaluar seguridad de los sistemas informáticos («Nmap: the Network Mapper - Free Security Scanner», s. f.)[4].

NMAP cuenta con varias funciones para analizar redes de computadoras, incluyendo detección de equipos, servicios y sistemas operativos. Resulta una herramienta muy útil para identificar posibles huecos en la seguridad y con ello corregir a tiempo dichos problemas.

Una tecnología que ayuda a mitigar las vulnerabilidades en los servicios de información web, son los certificados SSL ya que es un protocolo de seguridad que hace que los datos viajen de manera segura e íntegra, es decir, la transmisión de los datos entre el cliente y el servidor web es totalmente cifrada o encriptada, ya que emplean algoritmos matemáticos y sistema de claves que solo el cliente que tiene acceso al sistema web y el servidor pueden identificar.

Un certificado SSL implementa el modelo preferido de seguridad en web, contiene claves digitales que protegen la integridad de sus datos al momento de enviar y recibir. Los servidores que corren SSL crean una vía con un cifrado único para las sesiones privadas a través de Internet, la clave pública del servidor está al alcance de cualquier persona. Es por eso que utilizan una clave pública y una clave privada: La clave pública es para cifrar la información, la clave privada para descifrarla («¿Qué es un Certificado SSL? — Cert Superior», s. f.)[5].

Por último, la prueba de Turing automática y pública para diferenciar ordenadores de humanos ayudará a evitar que algoritmos que se ejecutan de forma automática en las aplicaciones web genere spam en los sistemas de información saturando las bases de datos con datos erróneos generados de forma independiente por dicho programa mal intencionado ya que al ser una prueba que ayuda a identificar si es un humano y/o persona real, el que está tratando de generar un registro en la base de datos. Típicamente esta prueba consiste en que una persona proporcione una serie de caracteres que se muestran de forma distorsionada en pantalla, de tal manera que solamente un humano pueda interpretarlas y no la máquina («¿ Que es un CAPCHA?

de Maria Betania Labus en Prezi», s. f.). La prueba de Turing a la que se refiere CAPTCHA es una prueba que Alan Turing propuso en 1950 para demostrar qué tanto la inteligencia de una máquina podría simular la de un humano, aun cuando superficialmente la diferencia fuera imperceptible.

Este método sigue siendo usado en inteligencia artificial. Como en CAPTCHA la prueba es controlada por una máquina, en realidad se está usando una prueba de Turing inversa. El uso de CAPTCHA existe en Internet desde 2003(«¿Qué es CAPTCHA?», s. f.)[6].

En base a los antecedentes de las tecnologías existentes que se menciona en la presente sección, se pretende desarrollar una propuesta de solución para la tesis descrita a continuación.

En la presente sección se muestran algunos de los conceptos utilizados en el desarrollo de este trabajo, los mismos que, ayudan a la comprensión del proyecto, y se presentan a continuación:

## 2.1. BackEnd

Define [31] a este como: las actividades realizadas del lado del servidor; es decir, las tareas de base de datos y los servidores de aplicaciones que el usuario no puede visualizar en el explorador de Internet. Los lenguajes usados comúnmente son PHP, Java, Ruby, .NET, ASP, Python, entre otros, los cuales son los encargados de interactuar con la base de datos”.

### 2.1.1. Php

Es un lenguaje de código abierto muy popular especialmente adecuado para el desarrollo web y que puede ser incrustado en HTML. En lugar de usar muchos comandos para mostrar HTML (como en C o en Perl), las páginas de PHP contienen HTML con código incrustado que hace .algo”(en este caso, mostrar ”¡Hola, soy un script de PHP!). El código de PHP está encerrado entre las etiquetas especiales de comienzo y final ¡?php y ?¿que permiten entrar y salir del ”modo PHP”. Lo que distingue a PHP de algo del lado del cliente como Javascript es que el código es

ejecutado en el servidor, generando HTML y enviándolo al cliente. El cliente recibirá el resultado de ejecutar el script, aunque no se sabrá el código subyacente que era. El servidor web puede ser configurado incluso para que procese todos los ficheros HTML con PHP, por lo que no hay manera de que los usuarios puedan saber qué se tiene debajo de la manga. Lo mejor de utilizar PHP es su extrema simplicidad para el principiante, pero a su vez ofrece muchas características avanzadas para los programadores profesionales. No sienta miedo de leer la larga lista de características de PHP. En unas pocas horas podrá empezar a escribir sus primeros scripts. Aunque el desarrollo de PHP está centrado en la programación de scripts del lado del servidor, se puede utilizar para muchas otras cosas. Siga leyendo y descubra más en la sección *¿Qué puede hacer PHP?*, o vaya directo al tutorial introductorio si solamente está interesado en programación web. («PHP: ¿Qué es PHP? - Manual», s. f.)<sup>[14]</sup>

### 2.1.2. MySQL

MySQL es la base de datos de código abierto más popular del mundo, lo que permite la entrega rentable de aplicaciones de base de datos integradas y confiables, de alto rendimiento y escalables. Es una base de datos integrada y segura para transacciones, compatible con ACID con capacidades completas de compromiso, retroacción, recuperación de fallas y bloqueo a nivel de fila. MySQL ofrece la facilidad de uso, la escalabilidad y el alto rendimiento, así como un conjunto completo de controladores de bases de datos y herramientas visuales para ayudar a los desarrolladores y administradores de bases de datos a crear y administrar sus aplicaciones de MySQL críticas para el negocio. MySQL es desarrollado, distribuido y respaldado por Oracle, y la información más reciente sobre el software MySQL se puede encontrar en el sitio web de MySQL. La base de datos MySQL proporciona las siguientes características:

- Alto rendimiento y escalabilidad para satisfacer las demandas de usuarios y cargas de datos en crecimiento exponencial.
- Clústeres de replicación autor regenerativos para mejorar la escalabilidad, el rendimiento y la disponibilidad.



- Cambio de esquema en línea para cumplir con los cambiantes requisitos comerciales.
- Esquema de rendimiento para supervisar el rendimiento a nivel de usuario y aplicación y el consumo de recursos.
- SQL y NoSQL Access para realizar consultas complejas y operaciones simples y rápidas de Key Value.
- Independencia de plataforma que le brinda la flexibilidad para desarrollar e implementar en múltiples sistemas operativos.
- Big Data Interoperability utilizando MySQL como el almacén de datos operativos para Hadoop y Cassandra. Para alcanzar el nivel más alto de escalabilidad, seguridad, confiabilidad y tiempo de actividad de MySQL , MySQL Enterprise Edition incluye el conjunto más completo de funciones avanzadas, herramientas de administración y soporte técnico, incluyendo MySQL Enterprise Monitor, MySQL Enterprise Backup, así como escalabilidad, seguridad y auditoría. y características de alta disponibilidad. MySQL Enterprise Edition también incluye Soporte Oracle Premier 24x7 y certificaciones e integraciones de productos Oracle.(«MySQL», s. f.)[20].

### 2.1.3. XML

Extensible Markup Language (XML) es un formato de texto simple y muy flexible derivado de SGML (ISO 8879). Originalmente diseñado para enfrentar los desafíos de la publicación electrónica a gran escala, XML también está desempeñando un papel cada vez más importante en el intercambio de una amplia variedad de datos en la Web y en otros lugares. («Google Traductor», s. f.-b)[24].

### 2.1.4. Base de Datos

En la obra [32] se afirma que: cada día, la mayoría de nosotros nos encontramos con actividades que requieren algún tipo de interacción con una base de datos (ingreso en un banco, reserva de una entrada para el teatro, solicitud de una suscripción a una

revista, compra de productos, ...). Estas interacciones son ejemplos de lo que se llama aplicaciones tradicionales de bases de datos (básicamente información numérica o de texto), aunque los avances tecnológicos han permitido que también existan: bases de datos multi-media, sistemas de información geográfica (GIS), almacenes de datos, sistemas de proceso analítico on-line”.

Una base de datos se entender a como una colección de datos relacionados entre sí y que tienen un significado implícito.

Por datos queremos decir hechos conocidos que pueden registrarse y que tienen un significado implícito.

## 2.2. FrontEnd

También según [31] “dentro del contexto del desarrollo de aplicaciones web, implica el uso de las tecnologías con las que interactúa directamente el usuario. Normalmente estas tecnologías son desarrolladas en los lenguajes de HTML, CSS y Javascript; también se usan las herramientas de diseño gráfico como Photoshop o Fireworks. El objetivo es desarrollar la interfaz gráfica de usuario (GUI), buscando una experiencia de uso bien valorada por el usuario final, siendo en algunos casos necesario hacer investigación, estudios y pruebas para llegar a esten. Además, dentro del desarrollo de las aplicaciones web es posible desarrollar el front-end de la aplicación sin contar con una aplicación back-end que interactúe con la base de datos”.

### 2.2.1. Dom

El Document Object Model (DOM) es una interfaz neutral de plataforma y plataforma que permitirá que los programas y scripts accedan y actualicen dinámicamente el contenido, la estructura y el estilo de los documentos. El documento se puede seguir procesando y los resultados de ese procesamiento se pueden incorporar nuevamente a la página presentada. Esta es una descripción general de los materiales relacionados con DOM aquí en W3C y en la web. ”HTML dinámico.<sup>es</sup> un término utilizado por algunos proveedores para describir la combinación de HTML, hojas de estilo y scripts que permiten animar documentos. El W3C ha recibido varias presentaciones de empresas miembros sobre la forma en que el modelo de objetos de documentos

HTML debe exponerse a las secuencias de comandos. Estas presentaciones no proponen ninguna nueva etiqueta HTML o tecnología de hoja de estilo. La actividad DOM del W3C está trabajando arduamente para garantizar que se acuerden soluciones neutrales interoperables y de lenguaje de scripting («Modelo de objetos de documento W3C», s. f.)[28].

### 2.2.2. Html

Significa "Lenguaje de Marcado de Hypertexto" por sus siglas en ingles "Hyper-Text Markup Language", es un lenguaje que pertenece a la familia de los "lenguajes de marcado" es utilizado para la elaboración de páginas web. El estándar HTML lo define la W3C (World Wide Web Consortium) y actualmente HTML se encuentra en su versión HTML5. Cabe destacar que HTML no es un lenguaje de programación ya que no cuenta con funciones aritméticas, variables o estructuras de control propias de los lenguajes de programación, por lo que HTML genera únicamente páginas web estáticas, sin embargo, HTML se puede usar en conjunto con diversos lenguajes de programación para la creación de páginas web dinámicas. («W3C HTML», s. f.)[15]

### 2.2.3. Diseño Web Adaptativo

Este término es usado en referencia a la manera en que cada explorador de Internet responde a su ambiente, aportando así a la mejora de la experiencia del usuario en el uso del sitio, independientemente del tamaño de resolución y pantalla del dispositivo. Para lograr esto es imprescindible el uso de CSS3 acompañado del uso media queries dentro del mismo CSS, [31].

Según [33] existen tres elementos fundamentales para lograr un diseño web adaptativo, que son:

- Cuadrícula fluida: Este concepto lo que propone es usar porcentajes para definir los tamaños de las columnas o divs, en lugar de píxeles.
- Imágenes flexibles: Las imágenes no tienen anchos fijos sino un máximo (max-width), que en una computadora de escritorio o laptop suele mostrarse al 100

- Media queries: Con el uso de estilos CSS permite que se personalice utilizando un ancho mínimo y máximo del navegador (min-max width).

#### 2.2.4. ECMAScript

Es una especificación de lenguaje de programación publicada por ECMA International. El desarrollo empezó en 1996 y estuvo basado en el popular lenguaje JavaScript propuesto como estándar por Netscape Communications Corporation. Actualmente está aceptado como el estándar ISO 16262.

ECMAScript define un lenguaje de tipos dinámicos ligeramente inspirado en Java y otros lenguajes del estilo de C. Soporta algunas características de la programación orientada a objetos mediante objetos basados en prototipos y pseudoclases.

La mayoría de navegadores de Internet incluyen una implementación del estándar ECMAScript, al igual que un acceso al Document Object Model para manipular páginas web. JavaScript está implementado en la mayoría de navegadores, Internet Explorer de Microsoft usa JScript. El navegador Opera tenía su propio intérprete de ECMAScript con extensiones para soportar algunas características de JavaScript y JScript, actualmente Opera está basado en Chromium (y utiliza su intérprete). Cada navegador tiene extensiones propias al estándar ECMAScript, pero cualquier código que se adecúe al estándar debería funcionar en todos ellos.

ActionScript, para Adobe Flash, también está basado en el estándar ECMAScript, con mejoras que permiten mover, crear y analizar dinámicamente objetos, mientras la película está en ejecución («Standard ECMA-262», s. f.)[27].

#### 2.2.5. Html5

HTML5 es un lenguaje markup (de hecho, las siglas de HTML significan Hyper Text Markup Language) usado para estructurar y presentar el contenido para la web. Es uno de los aspectos fundamentales para el funcionamiento de los sitios, pero no es el primero. Es de hecho la quinta revisión del estándar que fue creado en 1990. A fines del año pasado, la W3C la recomendó para transformarse en el estándar a ser usado en el desarrollo de proyectos venideros. Por así decirlo, qué es HTML5 está relacionado también con la entrada en decadencia del viejo estándar HTML 4, que

se combinaba con otros lenguajes para producir los sitios que podemos ver hoy en día. Con HTML5, tenemos otras posibilidades para explotar usando menos recursos. Con HTML5, también entra en desuso el formato XHTML, dado que ya no sería necesaria su implementación («HTML 5.2», s. f.)[22].

### 2.2.6. JavaScript

JavaScript es una importante tecnología web que provee de interactividad y efectos especiales las páginas web. Cada vez que una página web hace más que estar ahí y mostrar información estática para que puedas ver — mostrando actualizaciones de contenido oportunas, o mapas interactivos, o gráficos animados 2D/3D, o desplazarse por reproductores de video, etc. — puedes apostar a que JavaScript probablemente esté involucrado. JavaScript es posiblemente más difícil de aprender que las tecnologías relacionadas tales como HTML y CSS. Antes de intentar aprender JavaScript, se recomienda encarecidamente familiarizarse con al menos estas dos tecnologías primero, y quizás otras también. («JavaScript», s. f.)[17].

## 2.3. Entornos de desarrollo

### 2.3.1. Linux

Linux es un sistema operativo: un conjunto de programas que le permiten interactuar con su ordenador y ejecutar otros programas.

Un sistema operativo consiste en varios programas fundamentales que necesita el ordenador para poder comunicar y recibir instrucciones de los usuarios; tales como leer y escribir datos en el disco duro, cintas, e impresoras; controlar el uso de la memoria; y ejecutar otros programas. La parte más importante de un sistema operativo es el núcleo. En un sistema GNU/Linux, Linux es el núcleo. El resto del sistema consiste en otros programas, muchos de los cuales fueron escritos por o para el proyecto GNU. Dado que el núcleo de Linux en sí mismo no forma un sistema operativo funcional, preferimos utilizar el término “GNU/Linux” para referirnos a los sistemas que la mayor parte de las personas llaman de manera informal “Linux”.

Linux está modelado como un sistema operativo tipo Unix. Desde sus comienzos,

Linux se diseñó para que fuera un sistema multi tarea y multi usuario. Estos hechos son suficientes para diferenciar a Linux de otros sistemas operativos más conocidos. Sin embargo, Linux es más diferente de lo que pueda imaginar. Nadie es dueño de Linux, a diferencia de otros sistemas operativos. Gran parte de su desarrollo lo realizan voluntarios de forma altruista.

En 1984 comenzó el desarrollo de lo que más tarde sería GNU/Linux cuando la Free Software Foundation (Fundación de software libre, N. del t.) comenzó a desarrollar un sistema operativo libre de tipo Unix, llamado GNU.

El proyecto GNU ha desarrollado un conjunto de herramientas de software libre para ser utilizados por Unix<sup>TM</sup> y sistemas operativos tipo Unix como Linux. Estas herramientas permiten a los usuarios desarrollar tareas que van desde las mundanas (como copiar o eliminar ficheros del sistema) a las arcanas (como escribir y compilar programas o hacer edición sofisticada en una gran variedad de formatos de documento).

Aunque hay muchos grupos e individuos que han contribuido a Linux, la Free Software Foundation ha sido quien más ha contribuido. No sólo creó la mayor parte de las herramientas que se utilizan en Linux sino también la filosofía y comunidad que hizo que Linux fuera posible.

El núcleo Linux apareció por primera vez en 1991, cuando un estudiante de informática finlandés llamado Linus Torvalds anunció en el grupo de noticias de USENET comp.os.minix, una primera versión de un núcleo de reemplazo para Minix. Para más referencias consulte la página de historia de Linux en Linux Internacional.

Linus Torvalds sigue coordinando el trabajo de varios cientos de desarrolladores con la ayuda de cierto número de responsables de subsistemas. Existe una página oficial del núcleo Linux. Se puede encontrar un excelente resumen semanal de las discusiones en la lista de correo linux-kernel en Kernel Traffic. Puede encontrar más información sobre la lista de correo linux-kernel en el documento PUF de la lista de correo «linux-kernel».

Los usuarios de Linux tienen una gran libertad al elegir sus programas. Por ejemplo, un usuario de Linux puede elegir entre docenas de distintos intérpretes de línea de órdenes y entre distintos entornos de escritorio. Tantas opciones confunden a veces a los usuarios de otros sistemas operativos que no están acostumbrados a

poder modificar el intérprete de línea de órdenes o el entorno de escritorio.

Es menos probable que un sistema Linux se colapse, además tiene mejor capacidad para ejecutar múltiples programas al mismo tiempo y es más seguro que muchos otros sistemas operativos. Debido a estas ventajas, Linux es el sistema operativo que ha experimentado mayor crecimiento en el mercado de los servidores. Últimamente, Linux está empezando a ser popular entre los usuarios domésticos y en empresas («1.2. ¿Qué es GNU/Linux?», s. f.)[23].

### 2.3.2. Apache

El Proyecto Apache HTTP Server es un esfuerzo para desarrollar y mantener un servidor HTTP de código abierto para los sistemas operativos modernos, incluidos UNIX y Windows. El objetivo de este proyecto es proporcionar un servidor seguro, eficiente y extensible que proporcione servicios HTTP en sincronización con los estándares HTTP actuales. Apache HTTP Server ("httpd") se lanzó en 1995 y ha sido el servidor web más popular en Internet desde abril de 1996. En febrero de 2015, celebró su 20º aniversario como proyecto. El servidor Apache HTTP es un proyecto de The Apache Software Foundation («¡Bienvenido! - El proyecto del servidor Apache HTTP», s. f.)[25].

### 2.3.3. Sublime Text

Sublime Text es un editor de texto y editor de código fuente está escrito en C++ y Python para los plugin. Desarrollado originalmente como una extensión de Vim, con el tiempo fue creando una identidad propia, por esto aún conserva un modo de edición tipo vi llamado Vintage mode. Se puede descargar y evaluar de forma gratuita. Sin embargo, no es software libre o de código abierto<sup>3</sup> y se debe obtener una licencia para su uso continuado, aunque la versión de evaluación es plenamente funcional y no tiene fecha de caducidad. Características Minimapa: consiste en una previsualización de la estructura del código, es muy útil para desplazarse por el archivo cuando se conoce bien la estructura de este. Multi Selección: Hace una selección múltiple de un término por diferentes partes del archivo. Multi Cursor: Crea cursores con los que podemos escribir texto de forma arbitraria en diferentes posiciones del

archivo. Multi Layout: Trae siete configuraciones de plantilla podemos elegir editar en una sola ventana o hacer una división de hasta cuatro ventanas verticales o cuatro ventanas en cuadrícula. Soporte nativo para infinidad de lenguajes: Soporta de forma nativa 43 lenguajes de programación y texto plano. Syntax Highlight configurable: El remarcado de sintaxis es completamente configurable a través de archivos de configuración del usuario. Búsqueda Dinámica: Se puede hacer búsqueda de expresiones regulares o por archivos, proyectos, directorios, una conjunción de ellos o todo a la vez. Auto completado y marcado de llaves: Se puede ir a la llave que cierra o abre un bloque de una forma sencilla. Soporte de Snippets y Plugins: Los snippets son similares a las macros o los bundles además de la existencia de multitud de plugins. Configuración total de Keybindings: Todas las teclas pueden ser sobrescritas a nuestro gusto. Acceso rápido a línea o archivo: Se puede abrir un archivo utilizando el conjunto de teclas Cmd+P en Mac OS X o Ctrl+P en Windows y Linux y escribiendo el nombre del mismo o navegando por una lista. También se puede ir a una línea utilizando los dos puntos ":z el número de línea. Paleta de Comandos: Un intérprete de Python diseñado solo para el programa con el cual se puede realizar infinidad de tareas. Coloreado y envoltura de sintaxis: Si se escribe en un lenguaje de programación o marcado, resalta las expresiones propias de la sintaxis de ese lenguaje para facilitar su lectura. Pestañas: Se pueden abrir varios documentos y organizarlos en pestañas. Resaltado de paréntesis e indentación: Cuando el usuario coloca el cursor en un paréntesis, corchete o llave, resalta ésta y el paréntesis, corchete o llave de cierre o apertura correspondiente («Sublime Text 3 Beta ya disponible», s. f.)[18].

#### 2.3.4. Xampp

Es el entorno más popular de desarrollo con PHP XAMPP es una distribución de Apache completamente gratuita y fácil de instalar que contiene MariaDB, PHP y Perl. El paquete de instalación de XAMPP ha sido diseñado para ser increíblemente fácil de instalar y usar. Mucha gente conoce de primera mano que no es fácil instalar un servidor de web Apache y la tarea se complica si le añadimos MariaDB, PHP y Perl. El objetivo de XAMPP es crear una distribución fácil de instalar para desarrolladores que se están iniciando en el mundo de Apache. XAMPP viene configurado por defecto con todas las opciones activadas. XAMPP es gratuito tanto para



usos comerciales como no comerciales. En caso de usar XAMPP comercialmente, asegúrate de que cumples con las licencias de los productos incluidos en XAMPP. Actualmente XAMPP tiene instaladores para Windows, Linux y OS X («About the XAMPP project», s. f.)[26].

## 2.4. Intercambio de datos

### 2.4.1. Ajax

Es una técnica de desarrollo WEB, por la cual se pueden crear aplicaciones WEB más rápidas y cómodas para el usuario. Por medio de esta técnica el cliente puede interactuar con el servidor de manera asincrónica, actualizando las páginas, sin necesidad de volver a cargarlas.

Esta técnica, no solo es más cómoda y amigable para el usuario (ya que se asemeja a las aplicaciones de escritorio) sino que además es más rápida, porque cada vez que se necesita actualizar un dato en una página, no es necesario recargarla nuevamente (solo se recarga la sección necesaria de la misma).

AJAX, no es una tecnología en si, sino que es un conjunto de tecnologías aplicadas de manera que logran el resultado explicado anteriormente (es decir, logran AJAX). AJAX significa Asynchronous JavaScript And XML, y como su nombre lo indica, se trata de la combinación de JavaScript y XML. JavaScript hace las peticiones al servidor, el mismo le devuelve un resultado (response) en XML, y este es procesado por JavaScript para actualizar los datos de la página, sin tener que recargarla por completo (logrando así una interacción asincrónica entre el servidor y el cliente)(Just Sherekan Blog de Programación, Introducción a Ajax, 2008)[29].

### 2.4.2. JSON

(JavaScript Object Notation - Notación de Objetos de JavaScript) es un formato ligero de intercambio de datos. Leerlo y escribirlo es simple para humanos, mientras que para las máquinas es simple interpretarlo y generarlo. Está basado en un subconjunto del Lenguaje de Programación JavaScript, Standard ECMA-262 3rd Edición, diciembre 1999. JSON es un formato de texto que es completamente independiente

del lenguaje, pero utiliza convenciones que son ampliamente conocidos por los programadores de la familia de lenguajes C, incluyendo C, C++, C, Java, JavaScript, Perl, Python, y muchos otros. Estas propiedades hacen que JSON sea un lenguaje ideal para el intercambio de datos. JSON está constituido por dos estructuras: Una colección de pares de nombre/valor. En varios lenguajes esto es conocido como un objeto, registro, estructura, diccionario, tabla hash, lista de claves o un arreglo asociativo. Una lista ordenada de valores. En la mayoría de los lenguajes, esto se implementa como arreglos, vectores, listas o secuencias. Estas son estructuras universales; virtualmente todos los lenguajes de programación las soportan de una forma u otra. Es razonable que un formato de intercambio de datos que es independiente del lenguaje de programación se base en estas estructuras («JSON», s. f.)[21].

## 2.5. Seguridad informática

### 2.5.1. Hacker

Mientras que un hacker es simplemente un programador inteligente, experto en manipular o modificar un sistema o red informática, un hacker malicioso es alguien que utiliza sus conocimientos de informática para obtener acceso no autorizado a datos tales como información de tarjetas de crédito o imágenes personales, ya sea para diversión, beneficio, para causar daño o por otras razones. El hacking informático recurre a la manipulación de la conducta normal de un equipo y de los sistemas que tiene conectados. Esto se hace generalmente mediante scripts o programas que manipulan los datos que pasan a través de una conexión de red con el fin de acceder a la información del sistema. Las técnicas de hacking incluyen el uso de virus, gusanos, Troyanos, ransomware, secuestros del navegador, rootkits y ataques de denegación de servicio («Hacking (Hacker)», s. f.)[? ].

### 2.5.2. Spyware

Es un código malicioso cuyo principal objetivo es recoger información sobre las actividades de un usuario en un computador (tendencias de navegación), para permitir el despliegue sin autorización en ventanas emergentes de propaganda de mercadeo, o

para robar información personal (p.ej. números de tarjetas de crédito). Hay iniciativas de utilizarlos para controlar el uso de software pirata. Según algunas estadísticas, cerca del 91 % de los computadores tienen spyware instalado, y de acuerdo a un reporte de la firma EarthLink”, en una revisión de cerca de 1 millón de computadores en Internet, el promedio de programas “spyware” en cada uno era de 28[30].

### 2.5.3. Troyano

Son programas de código malicioso, que de diferentes maneras se alojan en los computadores con el propósito de permitir el acceso no autorizado a un atacante, o permitir el control de forma remota de los sistemas. El virus, adicionalmente, tiene como objetivo principal ser destructivo, dañando la información de la máquina, o generando el consumo de recursos de manera incontrolada para bloquear o negar servicios[30]. El vector de propagación de estos códigos es, casi siempre, otro programa o archivo (un programa ejecutable, imagen, video, música, reproducciones flash, etc.); de otra parte, los virus, se replican ellos mismos una vez instalados en el sistema. Las estadísticas indican que mensualmente se generan cientos de estos programas, cuyo principal objetivo es robo financiero, poniendo en riesgo la información confidencial y el dinero de las personas y de las organizaciones, más que la destrucción de archivos. La última tendencia en clases de virus se denomina cripto-virus, el cual, una vez instalado, cifra la información contenida en el disco del equipo, o algunos archivos contenidos en éste, y posteriormente se solicita una cantidad de dinero para que sus autores entreguen las claves para recuperar el contenido de los archivos cifrados (secuestro express de la información)[30].

### 2.5.4. Phishing

Es un ataque del tipo ingeniería social, cuyo objetivo principal es obtener de manera fraudulenta datos confidenciales de un usuario, especialmente financieros, aprovechando la confianza que éste tiene en los servicios tecnológicos, el desconocimiento de la forma en que operan y la oferta de servicios en algunos casos con pobres medidas de seguridad. Actualmente, los ataques de phishing son bastante sofisticados, utilizando mensajes de correo electrónico y falsos sitios Web, que suplantan

perfectamente a los sitios originales[30].

### 2.5.5. Botnets

Son máquinas infectadas y controladas remotamente, que se comportan como “zombis”, quedando incorporadas a redes distribuidas de computadores llamados robot, los cuales envían de forma masiva mensajes de correo “spam” o código malicioso, con el objetivo de atacar otros sistemas; se han detectado redes de más de 200.000 nodos enlazados y más de 10.000 formas diferentes de patrones de “bots”. Las organizaciones deberían revisar los computadores de sus redes de datos para detectar síntomas de infecciones relacionadas con este patrón, para evitar ser la fuente de ataques hacia otras redes o sistemas. También se requiere de la colaboración y aporte permanente de los usuarios finales y de los proveedores de acceso a Internet y prestadores de servicios como los “café Internet”[30].

### 2.5.6. Trashing

Un método cuyo nombre hace referencia al manejo de la basura. No es una técnica relacionada directamente con los sistemas de información, pues los atacantes se valen de otra forma de ingeniería social y para ello, el mecanismo utilizado, es la búsqueda en las canecas de la basura o en los sitios donde se desechan papeles y documentos de extractos bancarios, facturas, recibos, borradores de documentos, etc., y posteriormente utilizarla según convenga, elaborando un perfil de la víctima para robar su identidad, o teniendo acceso directamente a la información que se suponía confidencial[30].

## 2.6. Programación Orientada a Objetos (POO)

”La POO es un paradigma de programación (o técnica de programación) que utiliza objetos e interacciones en el diseño de un sistema”, dice [34]. También dice que tiene sus propias características y elementos que se mencionan a continuación: La POO está compuesta por una serie de elementos que se detallan a continuación.

- Clase: Una clase es un modelo que se utiliza para crear objetos que compartan un mismo comportamiento, estado e identidad.
- Objeto: Es una entidad provista de métodos o mensajes a los cuales responde (comportamiento); atributos con valores concretos (estado); y propiedades (identidad).
- Método: Es el algoritmo asociado a un objeto que indica la capacidad de lo que este puede hacer.
- Evento y Mensaje: Un evento es un suceso en el sistema mientras que un mensaje es la comunicación del suceso dirigida al objeto.
- Propiedades y atributos: Las propiedades y atributos, son variables que contienen datos asociados a un objeto.

### 2.6.1. Características conceptuales de la POO

La POO debe guardar ciertas características que la identifican y diferencian de otros paradigmas de programación. Dichas características se describen a continuación.

- Abstracción: Aislación de un elemento de su contexto. Define las características esenciales de un objeto
- Encapsulamiento: Reúne al mismo nivel de abstracción, a todos los elementos que puedan considerarse pertenecientes a una misma entidad.
- Modularidad: Característica que permite dividir una aplicación en varias partes más pequeñas (denominadas módulos), independientes unas de otras.
- Ocultación (aislamiento): Los objetos están aislados del exterior, protegiendo a sus propiedades para no ser modificadas por aquellos que no tengan derecho a acceder a las mismas.
- Polimorfismo: Es la capacidad que da a diferentes objetos, la posibilidad de contar con métodos, propiedades y atributos de igual nombre, sin que los de un objeto interfieran con el de otro.

- Herencia: Es la relación existente entre dos o más clases, donde una es la principal (madre) y otras son secundarias y dependen (heredan) de ellas (clases “hijas”), donde a la vez, los objetos heredan las características de los objetos de los cuales heredan.
  
- Recolección de basura: Es la técnica que consiste en destruir aquellos objetos cuando ya no son necesarios, liberándolos de la memoria.

## 2.7. Patrón de arquitectura de Software.

### 2.7.1. Modelo Vista Control (MVC)

La definición de [35] sobre el MVC es: “este patrón propone la separación en distintos componentes de la interfaz de usuario (vistas), el modelo de negocio y la lógica de control. Una vista es una ‘fotografía’ del modelo (o una parte del mismo) en un determinado momento. Un control recibe un evento disparado por el usuario a través de la interfaz, accede al modelo de manera adecuada a la acción realizada, y presenta en una nueva vista el resultado de dicha acción. Por su parte, el modelo consiste en el conjunto de objetos que modelan los procesos de negocio que se realizan a través del sistema. En una aplicación web, las vistas serían las páginas HTML que el usuario visualiza en el navegador. A través de estas páginas el usuario interactúa con la aplicación, enviando eventos al servidor a través de peticiones HTTP. En el servidor se encuentra el código de control para estos eventos, que en función del evento concreto actúa sobre el modelo convenientemente. Los resultados de la acción se devuelven al usuario en forma de página HTML mediante la respuesta HTTP”. (Bahit, 2011) describe de forma gráfica el funcionamiento básico del patrón MVC en la Figura 2.1.

## 2.8. Modelo de procesos del software

### 2.8.1. Modelo en espiral:

“Es un enfoque realista para el desarrollo de sistemas y de software a gran escala. Como el software evoluciona a medida que el proceso avanza, el desarrollador y cliente comprenden y reaccionan mejor ante los riesgos en cada nivel de evolución. El modelo espiral usa los prototipos como mecanismo de reducción de riesgos, pero, más importante, permite aplicar el enfoque de hacer prototipos en cualquier etapa de la evolución del producto. Mantiene el enfoque de escalón sistemático sugerido por el ciclo de vida clásico, pero lo incorpora en una estructura iterativa que refleja al mundo real en una forma más realista. El modelo espiral demanda una consideración directa de los riesgos técnicos en todas las etapas del proyecto y, si se aplica de manera apropiada, debe reducir los riesgos antes de que se vuelvan un problema” dice [36]. Se presenta en la Figura 2.1[39] el modelo en espiral del proceso de software propuesto por [37].

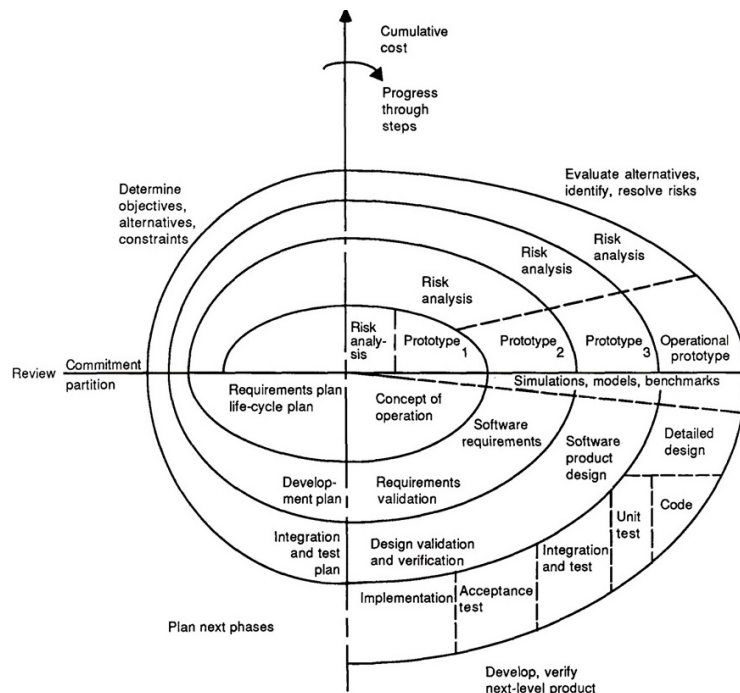


Figura 2.1: Modelo en espiral

## 2.9. Técnicas de aprendizaje automático

### 2.9.1. Algoritmos Genéticos (AGs)

son métodos adaptativos que pueden usarse para resolver problemas de búsqueda y optimización. Están basados en el proceso genético de los organismos vivos. A lo largo de las generaciones, las poblaciones evolucionan en la naturaleza de acorde con los principios de la selección natural y la supervivencia de los más fuertes, postulados por Darwin (1859). Por imitación de este proceso, los Algoritmos Genéticos son capaces de ir creando soluciones para problemas del mundo real. La evolución de dichas soluciones hacia valores óptimos del problema depende en buena medida de una adecuada codificación de las mismas.

Los principios básicos de los Algoritmos Genéticos fueron establecidos por Holland (1975), y se encuentran bien descritos en varios textos — Goldberg (1989), Davis (1991), Michalewicz (1992), Reeves (1993) — . En la naturaleza los individuos de una población compiten entre sí en la búsqueda de recursos tales como comida, agua y refugio. Incluso los miembros de una misma especie compiten a menudo en la búsqueda de un compañero. Aquellos individuos que tienen más éxito en sobrevivir y en atraer compañeros tienen mayor probabilidad de generar un gran número de descendientes. Por el contrario individuos poco dotados producirán un menor número de descendientes.

Esto significa que los genes de los individuos mejor adaptados se propagarán en sucesivas generaciones hacia un número de individuos creciente. La combinación de buenas características provenientes de diferentes ancestros, puede a veces producir descendientes "superindividuos", cuya adaptación es mucho mayor que la de cualquiera de sus ancestros. De esta manera, las especies evolucionan logrando unas características cada vez mejor adaptadas al entorno en el que viven.

Los Algoritmos Genéticos usan una analogía directa con el comportamiento natural. Trabajan con una población de individuos, cada uno de los cuales representa una solución factible a un problema dado. A cada individuo se le asigna un valor ó puntuación, relacionado con la bondad de dicha solución. En la naturaleza esto equivaldría al grado de efectividad de un organismo para competir por unos determinados recursos. Cuanto mayor sea la adaptación de un individuo al problema,



mayor será la probabilidad de que el mismo sea seleccionado para reproducirse, cruzando su material genético con otro individuo seleccionado de igual forma.

Este cruce producirá nuevos individuos descendientes de los anteriores los cuales comparten algunas de las características de sus padres. Cuanto menor sea la adaptación de un individuo, menor será la probabilidad de que dicho individuo sea seleccionado para la reproducción, y por tanto de que su material genético se propague en sucesivas generaciones [40].

```
BEGIN /* Algoritmo Genetico Simple */
  Generar una poblacion inicial.
  Computar la funcion de evaluacion de cada individuo.
  WHILE NOT Terminado DO
    BEGIN /* Producir nueva generacion */
      FOR Tamaño poblacion/2 DO
        BEGIN /*Ciclo Reproductivo */
          Seleccionar dos individuos de la anterior generacion,
          para el cruce (probabilidad de seleccion proporcional
          a la funcion de evaluacion del individuo).
          Cruzar con cierta probabilidad los dos
          individuos obteniendo dos descendientes.
          Mutar los dos descendientes con cierta probabilidad.
          Computar la funcion de evaluacion de los dos
          descendientes mutados.
          Insertar los dos descendientes mutados en la nueva generacion.
        END
      END
    END
  IF la poblacion ha convergido THEN
    Terminado := TRUE
  END
END
```

Figura 2.2: Pseudo código del Algoritmo Genético Simple

---

## Capítulo 3

### Estado del Arte

En términos generales, la seguridad de los sistemas de información orientados a la web, debe ser primordial ya que los ciberataques están a la orden del día, desde una denegación de servicio, comúnmente llamado DDOS, puede dejar inhabilitado el acceso al sistema, ocasionando que cualquier tipo de usuarios no pueda consultar su información. Como el ejemplo anterior se pueden mencionar muchos más, de acuerdo a lo antes mencionado, se estructura el presente estado del arte. (MSc. Susana C. Romaniz, s. f.) Menciona que la seguridad de las aplicaciones web es un tema de interés actual y cotidiano ya que las complejas y sensibles funciones de los sistemas web han movido el perímetro de seguridad de las organizaciones. Un simple descuido puede dejar desprotegidas funciones críticas de dichos sistemas web [8].

Se dice que los nuevos ataques esta concebido de forma tal que resulta imposible considerarlos al momento del desarrollo de las aplicaciones. Algunos problemas han perdido importancia a medida que se han incrementado la concientización respecto a ellos y algunas categorías de defectos se han eliminado como consecuencia de los cambios introducidos en el software del navegador web. La dura batalla por proteger los sistemas web es probable que siga así en el futuro mediato. Un ejemplo “Este sitio es absolutamente seguro. Ha sido diseñado para utilizar tecnología 128-bit SSL a fin de evitar que usuarios no autorizados visualicen su información. Ud. puede utilizar este sitio con la tranquilidad de que sus datos están seguros con nosotros.”. La mayoría de las aplicaciones web sin inseguras debido a condiciones que no tienen ninguna relación con SSL. Se mencionan los ataques más comunes: i) Quiebre de la autenticación, ii) Quiebre de los controles de acceso, iii) Inyección de SQL, iv) Cross-site scripting, v) Fuga de información.

(Ortega Martorell y Canino Gutiérrez, 2006) explica que el uso creciente de internet, ha dirigido la atención hacia un problema crucial: la privacidad. Para que tenga un lugar una negociación en internet, se precisa, en la mayoría de los casos, que cada entidad participante pueda contar con una manera eficaz de verificar la identidad de otras y establecer un nivel de confianza, Es necesario, por tanto, crear un entorno que garantice la autenticidad y seguridad de las transacciones que tienen lugar en este proceso. SSL (Secure Sockets Layer) es el protocolo de seguridad mas estandarizado que, haciendo uso de un conjunto de técnicas criptográficas, asegura confidencialidad e integridad de la información[9].

SSL es un protocolo de propósito general para establecer comunicaciones seguras. No es exclusivo del comercio electrónico, sino que sirva para cualquier comunicación vía internet, y por lo tanto también para las transacciones económicas.

(Muro y Ramírez, 2000) describe que el uso de las redes de computadoras está presente en casi todas las actividades humanas, tales como la educación, las transacciones bancarias, las declaraciones de impuestos, y en especial en el comercio, entre otras[10]. Sin embargo, a pesar de su amplio uso, las redes aún enfrentan serios problemas en cuestiones de seguridad. La mayoría de los problemas de seguridad son causados intencionalmente por gente que intenta ganar algo o hacerle daño a alguien. En este artículo menciona algunos ejemplos de personas malintencionadas que buscan obtener provecho de las vulnerabilidades de los sistemas web, a continuación, se detallan:

i) Estudiante: Divertirse husmeando el correo de otra gente, ii) Hacker: Probar el sistema de seguridad de alguien; robar datos, iii) Representante de ventas: Indicar que representa a toda Europa, no sólo a Andorra, iv) Hombre de negocios: Descubrir el plan estratégico. Ex empleado: Vengar su despido, v) Contador: Estafar dinero de una compañía, vi) Corredor de bolsa: Negar una promesa hecha a un cliente por correo electrónico, vii) Timador: Robar números de tarjeta de crédito, viii) Espía: Conocer la fuerza militar de un enemigo, ix) Terrorista: Robar secretos de guerra bacteriológica.

SSL fue inicialmente concebido como un protocolo para la seguridad Web, pero actualmente es una nueva capa de protocolos que opera sobre el protocolo TCP de Internet. Se puede usar para proteger la comunicación de cualquier protocolo de aplicación que opere normalmente sobre TCP, por ejemplo, HTTP, FTP, ó TELNET. El uso más común de SSL consiste en proteger las comunicaciones HTTP: en particular, una URL que comienza con “https://” indica el uso de HTTP bajo la protección de SSL.

(Muro y Ramírez, 2000) Constata que las redes digitales de comunicación son inherentemente inseguras y que deben ser provistas, necesariamente, de mecanismos que permitan a los usuarios utilizarlas asumiendo el menor riesgo posible. Esta condición necesaria, que a veces no es suficiente, es aún más exigible cuando se trata de realizar comercio a través de ellas o se transfieren datos personales de carácter

inherentemente confidencial[10].

(Vega y Vinasco-Salazar, 2014) Describe los CAPTCHA como las pruebas humanas interactivas más utilizadas, aparecen como un método de diferenciación entre usuarios humanos y máquinas para brindar seguridad a la información en internet y evitar el spam, especialmente[11]. Son similares a la prueba de Turing, con la diferencia de que el juez es un equipo cuyo objetivo es hacer preguntas que los usuarios humanos podrán responder, pero las máquinas actuales no. Otra descripción de los CAPTCHA, son unos "puzzles" que los webmasters incluyen en su sitio web para asegurarse de que los visitantes que quieren interactuar con el contenido son personas, y no robots spam que tratan de registrarse en el sitio web, incluir comentarios en blogs, etc. Concluye diciendo que no hay desacuerdo alguno sobre la imperante necesidad de proporcionar seguridad a la información personal e institucional colocada en internet, que se incrementa exponencialmente día a día, ante la gran cantidad y constante posibilidad de ataques informáticos. Sin embargo, con soluciones como los CAPTCHA se han logrado resultados en relación con esto.

(López y Rodríguez, 2016) abordan la seguridad que existe actualmente en internet, las principales amenazas que se pueden sufrir a través de este medio, así como también que medidas de seguridad se pueden implementar. Se dice que en décadas pasadas los programadores desarrollaron códigos maliciosos que ayudan al robo de información, ya que anteriormente el número de páginas no seguras era mayor, no existía la encriptación de los datos que viajan en la red para proteger dicha información[12]. Para que esto se desarrollaron algunos métodos con la finalidad de mejorar la seguridad en la comunicación y transacciones que se realizan diariamente, Algunos métodos criptográficos son: i) Simétricos: este método hace uso de una clave, que se utiliza tanto para ocultar el mensaje y como para descubrirlo, ii) Asimétricos: consta de dos claves, una pública que todas las personas conocen y una privada que sólo las personas que envían y reciben la conocen. Para encriptar se realiza una clave pública y para el proceso inverso es necesaria la clave privada del destinatario.

Estos métodos aseguran que el mensaje no sea comprensible para cualquiera que no tenga la llave, sin embargo, no garantizan que el medio por el que se transmiten sea seguro, por lo que se apoyan en sistemas de encriptación para que el mensaje llegue al destinatario por un medio seguro. Para lo anterior se hace uso de sistemas

de encriptación como lo son el SSL, protocolo que asegura el transporte de información a través de la red, y el Secure Hypertext Transfer Protocol (SHTTP), variante del SSL, brinda mayor seguridad en la transferencia de información por un medio inseguro, internet. Con lo anterior así mitigar el robo de información.

(López y Rodríguez, 2016) Concluyen “No debemos olvidar que, como usuarios, también tenemos la responsabilidad de asegurarnos que nuestra información llegue a su destino y que serán utilizadas de manera correcta, o para el fin que fueron dados, razón por la cual debemos de contar con buenas prácticas, antivirus, mantener actualizados el software que utilicemos, etcétera, mientras navegamos en internet y con esto asegurarnos que todo lo que enviamos no pueda ser interceptado o contar con una buena reputación en internet” [12].

(Nodarse, 2013) Hacen especial hincapié en que la seguridad en el desarrollo de los sistemas web depende de tres factores esenciales como son la tecnología, las organizaciones y el hombre, y la corrección de dichos problemas de seguridad que pudieran surgir, recaen en ellos. La seguridad en las organizaciones debe ser un problema visto por todos, ya que es un proceso constante de mejora el cual debe de evolucionar y adaptarse[13].

La seguridad es un aspecto crítico del comercio electrónico y la mejora en estos procesos podría expandir la confianza de los clientes. Se estima que a medida que la distancia entre procesos de negocio y tecnologías disminuye, el impacto de riesgos de seguridad aumenta(Nodarse, 2013). (Nodarse, 2013) Propone que las políticas de seguridad se establecen y se desarrollan aplicando los siguientes requisitos mínimos:

i) Organización e implantación del proceso de seguridad, ii) Análisis y gestión de los riesgos, iii) Gestión de personal, iv) Profesionalidad, v) Autorización y control de los accesos, vi) Protección de las instalaciones, vii) Adquisición de productos de seguridad, viii) Seguridad por defecto, ix) Integridad y actualización del sistema, x) Protección de la información almacenada y en tránsito, xi) Prevención ante otros sistemas de información interconectados, xii) Registro de actividad, xiii) Incidentes de seguridad, xiv) Continuidad de la actividad, xv) Mejora continua del proceso de seguridad.

La seguridad es un proceso continuo, iterativo e incremental de mejora, que garantice su evolución y adaptación a fin de reducir las vulnerabilidades que se pudieran presentar y ocasionar problemas a las organizaciones y/o usuarios. Una de las medidas adoptadas por la computación en la nube es la implementación y uso de HTTPS/SSL para una navegación segura. La anterior es una medida muy eficaz contra el robo de información o intrusión ya que proporciona un encriptado entre el cliente y el servidor que imposibilita que los datos sean incomprensibles por terceros o en este caso ciber delincuentes o ciber atacantes.

(Vega 2008) dice que la seguridad significa disponer de medios que permitan reducir lo más que se pueda, la vulnerabilidad de la información y de los recursos; aunque no se puede alcanzar el 100 por ciento de seguridad, la tendencia debe ser llegar a ese valor extremo. Los hackers, crackers están vigilando permanentemente las redes con el fin de encontrar las vulnerabilidades o debilidades de un sistema de información, el desarrollo del software ha permitido hacer cada vez más fácil la configuración y su utilización, Internet también permite la conectividad de todo tipo de usuario, de esta forma las amenazas a la seguridad de la Información están latentes y en cualquier momento un servidor o dispositivo de red puede ser atacado con fines negativos a la imagen de la Empresa o Institución, a su funcionalidad y otros aspectos

(H. Tarazona, 2010) menciona que según los datos de la encuesta anual de Seguridad del FBI, los virus informáticos siguen siendo la principal fuente de pérdida financiera en las organizaciones, seguidos por los impactos derivados de accesos no autorizados a los sistemas, el robo de información de propiedad industrial, y la pérdida de computadores personales o elementos de computación móvil. Estas causas generan más del 74% del total de las pérdidas financieras. En conclusión, para en el presente estado del arte existen varias tecnologías que ayudan a mitigar las vulnerabilidades de los sistemas de información web, por resaltar algunos, se puede mencionar la implementación del protocolo https con ayuda de un certificado SSL para que la información que se envía entre el canal de comunicación que existe entre el cliente y el servidor viaje cifrada/encriptada, y de esta forma no sea comprensible para terceros (ciber atacantes, hackers etc.).

Con lo anterior se ayuda a que la información viaje de forma íntegra y privada. A su vez, también la implementación de la prueba de Turing completamente automática

y pública para diferenciar ordenadores de humanos o CAPTCHA por sus siglas en inglés también son una herramienta útil para evitar que algoritmos mal intencionados generen spam en formularios de registro en los sistemas de información, por ello saturan las bases de datos, es por eso que con la ayuda de esta tecnología se evita este problema y otros.

También con la ayuda de herramientas de software de análisis de seguridad, se pueden detectar una gran cantidad de vulnerabilidades, mismas que pudiera presentar el sistema de información web. Una vez detectadas, se procede a solucionar dichas vulnerabilidades y con la utilización de programas de pruebas de penetración, verificar que estos problemas son solucionados de una manera eficiente sin comprometer el funcionamiento del sistema.

Para finalizar, el estado de arte que aquí se presenta, amplió el panorama sobre las tecnologías que se van utilizar en el desarrollo de presente tesis, así como sus ventajas, desventajas y distintas implementaciones. Gracias a esto, se obtuvo un conocimiento más concreto y sustentado para desempeñar un buen trabajo en la problemática a solucionar.



---

## Capítulo 4

### Metodología

En este proyecto se utilizará al paradigma de la programación orientada a objetos y se aplicará el análisis y diseño en espiral, especialmente por la capacidad de adaptación gradual de este enfoque de desarrollo, con esto se gana en flexibilidad, interacción y un mayor ajuste a las necesidades del cliente, mismas que pudieran surgir en el desarrollo del proyecto, para una mayor comprensión ver Figura 4.1 modelo en espiral[36].

Mediante la reutilización y el uso de la herencia reducir al mínimo los códigos generadores de la interfaz de usuario y también la transferencia de la información entre los clientes y el servidor.

Un cliente web debe estar basado o extender un componente plantilla adaptativo.

Los principios de la Programación Orientada a Objetos deben prevalecer a lo largo de todo el desarrollo con la finalidad de promover la creación de componentes de software y la reutilización.

Todas las herramientas de desarrollo han de ser OpenSource en categoría de estables.

La comunicación entre los clientes y el servidor siempre será asíncrona y solo será posible mediante el listener del módulo BackEnd, el cual arbitra el acceso al servidor.

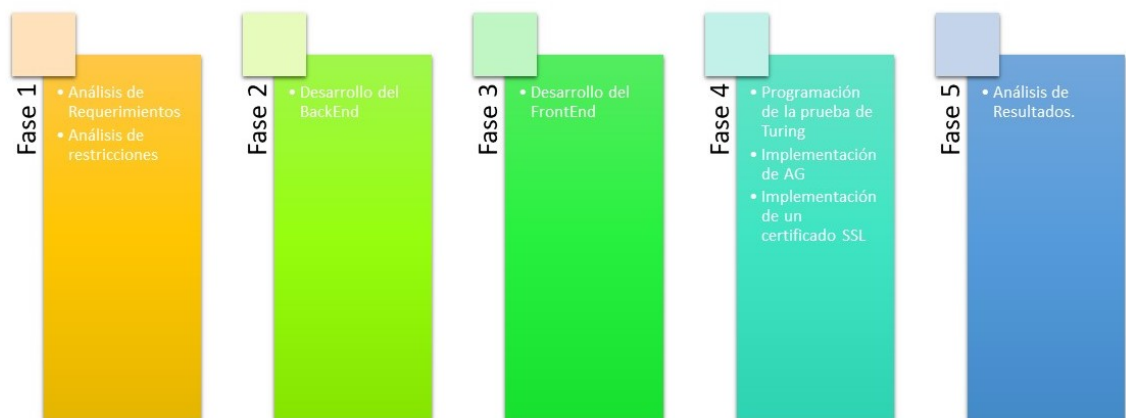


Figura 4.1: Diseño gráfico de la metodología.

*Fuente: Elaboración propia, FAHG*

## 4.1. Algoritmo Genético

Los algoritmos genéticos que simulan la evolución utilizando el código de computadora para ayudar a encontrar soluciones óptimas cuando se enfrentan problemas que involucran múltiples requisitos, son parte de una clase de algoritmos evolutivos que conforman un segmento más grande, el software de aprendizaje automático.

Como son la integración de una prueba de turing con algoritmos genéticos (AG), para la identificación entre un humano y un algoritmo mal intencionado (spam), que busca el acceso a un sistema de información mediante la autenticación a través de un formulario.

Los algoritmos evolutivos tienen su aplicación en la ingeniería en problemas complejos de optimización. De manera particular en esta investigación se utiliza un AG con la estructura de una prueba de turing, con el fin de diferenciar computadoras de humanos.

El funcionamiento del algoritmo genético parte de una población inicial, es definida por un conjunto de individuos. Cada individuo de la población es una propuesta de solución al problema. En este caso de estudio, la población inicial se definió aleatoriamente. Cada individuo es evaluado mediante una función objetivo con el propósito de evaluar la aptitud de este, para después pasar al proceso de selección.

Una vez seleccionado un par de individuos, se cruzan para generar dos nuevos individuos. Tras el cruce se da lugar a la mutación, consiste en modificar a uno o varios de sus genes de manera aleatoria. Para el criterio de convergencia o terminación, en este caso se limitó a un número máximo de evoluciones (generaciones).

El pseudocódigo de un algoritmo evolutivo, se expresa de la siguiente forma:

```
BEGIN / * Algoritmo evolutivo Simple * /  
Generar una población inicial  
Computar la función de evaluación de cada individuo.  
FOR Generaciones DO  
BEGIN / * Producir nueva generación * /  
FOR Sobrevivientes DO  
BEGIN
```

```
Seleccionar aleatoriamente individuo sobreviviente.  
Mutar sobreviviente con cierta probabilidad.  
Insertar sobreviviente en la nueva población.  
END  
FOR Poblaciontotal-Sobrevivientes DO  
BEGIN / * Ciclo Reproductivo * /  
Seleccionar dos individuos de la anterior  
generación, para el cruce (Selección por ruleta).  
Cruzar con cierta probabilidad los dos individuos  
obteniendo dos descendientes.  
Mutar descendiente con cierta probabilidad.  
Computar la función de evaluación del descendiente mutado.  
Insertar descendiente mutado en la nueva generación.  
  
END  
END  
END
```

#### 4.1.1. Codificación de los genes

El primer paso es determinar cómo representar los datos, esta es la base de la AG, que representa al individuo a través de genes que imitan algún elemento de los datos, para esto se representaron los genes como caracteres (caracteres ASCII), cada carácter puede provenir de uno de estas 26 letras: ABCDEFGHIJKLMNOPQRSTUUVWXY. A su vez se necesita un objetivo, en este caso, es la palabra que el usuario ingrese.

#### 4.1.2. Población inicial

La población inicial parte de la inicialización del individuo que consiste en pre-assignar los valores de los caracteres en la codificación de los genes, es principalmente una clase de conveniencia que mantiene a todos nuestros individuos unidos como una

unidad.

### 4.1.3. Selección

para garantizar que los mejores individuos tienen una mayor posibilidad de ser padres, frente a los individuos menos buenos. se debe de tener precaución para dar una oportunidad de reproducirse a los individuos menos buenos. Éstos pueden incluir material genético útil en el proceso de reproducción. Esta idea define la presión selectiva que determina en qué grado la reproducción está dirigida por los mejores individuos. En la clase se usa un enfoque de selección de ruleta (Seleccionar por genes aleatorios y elegir el mejor).

### 4.1.4. Cruza

Se puede tener uno o más operadores de cruce para la representación. Algunos aspectos importantes a tener en cuenta son:

- Los hijos deberían heredar algunas características de cada padre.
- Si éste no es el caso, entonces estamos ante un operador de mutación.
- Se debe diseñar de acuerdo a la representación.
- La recombinación debe producir cromosomas válidos.
- Se utiliza con una probabilidad alta de actuación sobre cada pareja de padres a cruzar, el punto de cruce puede ser entre 0.5 y 0.9, si no actúa los padres son los descendientes del proceso de recombinación de la pareja.

Para el experimento descrito en el presente el punto de cruce (tasa uniforme) fue determinado al 0.5 %.

### 4.1.5. Mutación

Al igual que en el cruce, en la mutación se puede tener uno o más operadores de mutación para nuestra representación. Pero existen aspectos a considerar para realizar el proceso:

- Debe permitir alcanzar cualquier parte del espacio de búsqueda.
- El tamaño de la mutación debe ser controlado.
- Debe producir cromosomas válidos.
- Se aplica con una probabilidad muy baja de actuación sobre cada descendiente obtenido tras aplicar el operador de cruce.

La tasa de valores aleatorios está mutada al 0.01 %.

### 4.1.6. Resumen de parametros del AG

Cuadro 4.1: Resumen de los parámetros del algoritmo evolutivo.

<b>Parámetro</b>	<b>Valor</b>
Tamaño de la población.	Depende del objetivo.
Tamaño del Cromosoma	27
Selección	Ruleta
Porcentaje de cruce	0.5 %
Probabilidad de mutación	0.1 %
Numero de generaciones	200
Corridas	6
Métrica	Menor Costo

El ambiente de desarrollo del algoritmo evolutivo se realizó en un lenguaje de programación de PHP. La funcionalidad del algoritmo evolutivo se rige mediante los parámetros descritos en el Cuadro 5.1.

---

## Capítulo 5

### Resultados y Análisis

## **5.1. Descripción de la Interfaz**

En la presente sección, se describen los resultados obtenidos de la investigación, como son la integración de una prueba de Turing con algoritmos genéticos para la identificación entre un humano y un algoritmo mal intencionado(spam), el cual busca el acceso a un sistema de información mediante el logeo a través de un formulario. los Algoritmos Genéticos (AG) que simulan la evolución utilizando el código de computadora para ayudar a encontrar soluciones óptimas cuando se enfrentan problemas que involucran múltiples requisitos. son parte de una clase de algoritmos evolutivos que conforman un segmento más grande, el software de aprendizaje automático.



En la Figura 5.1 Integración AG, muestra la composición de un algoritmo genético y una prueba de Turing, la cual le solicita al usuario la siguiente instrucción: ¿Eres un humano? Forma una palabra con los siguientes caracteres: (ABCDEFGHIJKLMNOSTUVWXYZ).

The screenshot shows a web interface for system access. At the top, there are logos for SEP, the National Technological Institute of Mexico, and the Superior Technological Institute of Misantla. The main heading is 'Acceso al sistema'. Below this, there is a form with two input fields: 'Usuario:' containing 'XXXX0019HZGRR00' and 'Contraseña:' with masked characters. Below the password field is a Turing test question: '¿Eres un humano? Forma una palabra con los siguientes caracteres: (ABCDEFGHIJKLMNOSTUVWXYZ)'. A 'Validar' button is positioned below the question. At the bottom of the page, there is a small footer: '2018 - Derechos reservados - Tecnológico Nacional de México'.

Figura 5.1: Integración AG.

*Fuente: Elaboración propia, FAHG*

Posteriormente, se ingresa la palabra que el usuario decida siempre y cuando sea en mayúsculas y se elabore con las letras que le muestra en pantalla, se da clic en validar, para verificar que el humano está cumpliendo con la instrucción y no sea un algoritmo mal intencionado que requiere saturar el sistema con la solicitud de peticiones, como se puede observar en la , Figura 5.2.. Ingreso de la palabra en AG.

This screenshot is identical to Figure 5.1, but the text area below the Turing test question now contains the word 'TURING' in all caps. The 'Validar' button remains below the text area. The rest of the interface, including the user and password fields and the top navigation, is the same as in Figure 5.1.

Figura 5.2: Ingreso de la palabra en AG.

*Fuente: Elaboración propia, FAHG*

Después de dar clic en la validación, se corre el algoritmo genético, el cual internamente tiene la regla que se le mostró al usuario, que le indica que formule una palabra con los siguientes caracteres “ABDCDEFGHIJKLMNOPQRSTUVWXYZ”,

el AG vuelve la palabra tecleada por el usuario su objetivo, con base en esto debe de lograr la palabra ingresada a partir de dicha restricción, cabe mencionar que la palabra puede ser cualquiera, con cualquier carácter conocido, pero para tales casos donde no se respete la norma no se alcanzará la meta, siempre y cuando se cumpla el criterio el AG llegará a su objetivo y se habilitará el botón desbloquear, como se puede observar en la Figura 5.3. Corrida AG..



Figura 5.3: Corrida del AG.

*Fuente: Elaboración propia, FAHG*

La Figura 5.4, Botón ingresar, se puede observar que después de dar clic en el botón desbloquear, si el AG llegó a su objetivo, se habilita el botón de ingresar y se accede a la pantalla principal.



Figura 5.4: Se habilita en botón ingresar.

*Fuente: Elaboración propia, FAHG*

En la Figura 5.5, Autenticación correcta, se puede visualizar la pantalla del sistema, después de cumplir los requisitos previos para el mismo.



Figura 5.5: Logeo satisfactorio.  
*Fuente: Elaboración propia, FAHG*

En caso de que el usuario no comprenda la instrucción eh ingrese algo erróneo o que el algoritmo genético no alcance el objetivo, el sistema enviará un mensaje de error, como se puede observar en la Figura 5.6. Error AG, posterior a dar clic en desbloquear, ya que es la parte del proceso donde se valida que el AG llegue al objetivo (la palabra que ingresa el usuario final).



Figura 5.6: Error AG.  
*Fuente: Elaboración propia, FAHG*

La instrucción que se le muestra al usuario en pantalla es la siguiente: ¿Eres un humano? Forma una palabra con los siguientes caracteres: (ABCDEFGHIJKLMN OPQRSTUVWXYZ). El algoritmo genético que desarrollado en el lenguaje de programación php, ya que es un lenguaje de programación que se ejecuta del lado del servidor, a su vez se evita que personas mal intencionadas pudiera intentar vulnerar dicho código. Tambien se concidero que es un de codigo abierto, por lo que no requiere licencias. La integración se fue echa con HTML5, Javascript, DOM, JSON y PHP.

## 5.2. Descripción del código

En la sección anterior se pudo apreciar cada uno de los componentes con los que cuenta la interfaz, así como también las funciones que realizan dichos objetos, se describió paso a paso el proceso a realizar y el uso adecuado del mismo, para que el usuario final tenga conocimiento del mismo. A continuación se presenta el codigo desarrollado para que la interfaz funcione.

```
<!DOCTYPE html>
<html lang="es">
<head>
<meta charset="UTF-8">
<meta name="viewport" content="width=device-width, initial-scale=1.0">
<link rel="stylesheet" href="plantilla/css/mensajes.css">
<link rel="stylesheet" href="login/css/index.css">
<script src="/sitm/jquery.min.js"> </script>
<script src="/sitm/ie_eventsource.js"></script>
<script>
//INICIO FUNCION PARA DIBUJAR ELEMENTOS
function DrawBotBoot()
{
document.write("Forma una palabra con los siguientes caracteres:");
document.write("<br>");
```

```
document.write("(ABCDEFGHJKLMNOPQRSTUVWXYZ)");
document.write("<input id='BotBootInput' type='text'
  maxlength='8' size='8'/>");
document.write("<form id='form1' ><input id='btnRun' type='submit'
  value='Validar' /><br><div contenteditable='true'
  class='large_text_input' id='my_solution' value='Respuesta'
  style='background-color:#D3D1C1; border-bottom: 1px solid red' ></div>
  </br></form><span id='best_individual' class='large_text'>
  </span> <span id='message'> </span>");
document.getElementById('my_solution').style.visibility="hidden";
document.getElementById('best_individual').style.visibility="hidden";
}
//FIN FUNCION PARA DIBUJAR ELEMENTOS
//INICIO FUNCION PARA ESTABLECER LA
//CONEXION CON LAS FUNCIONES DEL ALGORITMO GENETICO
function run_server_ga(solution)
{

console.log("Calling Server looking for JSON in "+solution);

if (!!window.EventSource) {

var source= new EventSource("ga_sse_server.php?solution="+solution);

source.addEventListener('update', function(e)
{
console.log("Receiving JSON server side events");
var data = JSON.parse(e.data);
$('#best_individual').html(data.best_individual);
$('#best_fittest_value').text(data.best_fittest_value); //
$('#generation').text(data.generation);
$('#stagnant').text(data.stagnant);
```

```
$('#max_fitness').text(data.max_fitness);
$('#message').html(data.message);
$('#elapsed').text(data.elapsed);

if (data.done==true)

source.close();

}, false);
source.onerror = function(e) {
$('#message').html("EventSource failed.");

};
} else {
$('#message').html("<strong>Sorry your Browser doesn't support
SERVER SIDE Events , needed to stream the live results.
</strong>-<br>Supported browsers see here: <a href='http://caniuse.com/#feat=even
}
}
$(document).ready(function() {
console.log("document.ready");
// When any <submit > button is clicked
$('#btnRun').click( function() {
document.getElementById('my_solution').
innerText=document.getElementById('BotBootInput').value;
var solution = $("#my_solution").text();
run_server_ga(solution); //lets go to the server and look for this string
//SELECCIONA EL ELEMENTO QUE SERA REPLAZADO
var el = document.getElementById('btnRun');
//CREA UN NUEVO ELEMENTO QUE SERA ASIGNADO A el
var newEl = document.createElement('p');
// REPLAZA EL NUEVO ELEMENTO CON newEL
```

```
el.parentNode.replaceChild(newEl, el);
newEl.innerHTML = "<input id='Button1' type='button' value='Desbloquear' onclick=
return false; // EVITA QUE LA PAGINA NO SE REFRESQUE
ValidBotBoot();
}); //FIN DEL EVENTO HANDLER
}); //end document.ready
function ValidBotBoot(){
var d = document.getElementById('my_solution').textContent;
var x = document.getElementById('best_individual').textContent;
if (d == x) {
document.getElementById('Button1').style.visibility="hidden";
document.getElementById('btnIngresar').style.visibility="visible";
return true;
}
else if (d !=x ){
new mensajes().alerta("¡Error!, codigo incorrecto, por favor verifique");
document.getElementById('my_solution').textContent="";
document.getElementById('BotBootInput').value="";
var el = document.getElementById('Button1');
var newEl = document.createElement('p');
el.parentNode.replaceChild(newEl, el);
newEl.innerHTML = "<input action='#' id='btnRun' type='button' value='Validar' o
// PRUEBAS DE EJECUCION
$('#btnRun').click( function() {
document.getElementById('my_solution').
innerText=document.getElementById('BotBootInput').value;
var solution = $("#my_solution").text();
run_server_ga(solution);
//SELECCIONA EL ELEMENTO QUE SERA REPLAZADO
var el = document.getElementById('btnRun');
//CREA UN NUEVO ELEMENTO QUE SERA ASIGNADO A el
var newEl = document.createElement('p');
```

```
// REPLAZA EL NUEVO ELEMENTO CON newEL
el.parentNode.replaceChild(newEl, el);
newEl.innerHTML = "<input id='Button1' type='button' value='Desbloquear' onclick=
return false; //
ValidBotBoot();
});
}
}
</script>
<!-- TERMINA CODIGO PARA ALGORITMO GENETICO -->
<!-- INICIA LA SEMANTICA DE LA PAGINA -->
<script src="plantilla/js/header.js"></script>
<script src="plantilla/js/footer.js"></script>
<script src="plantilla/js/mensajes.js"></script>
<script src="plantilla/js/validar.js"></script>
<script src="plantilla/js/gdata.js"></script>
<script src="plantilla/js/Ajax.js"></script>
<script src="plantilla/js/utils.js"></script>
<script src="login/js/index.js"></script>

<title>TecNM - Acceso</title>
</head>
<body>
<section id="main">
<article id="login">
<fieldset>
<div id="pUsuario"><label>Usuario: <input type="text" id="usuario" value="FIGJ701
<div id="pPwd"><label>Contraseña: <input type="password" id="pwd" value="12345"><
</fieldset>
<fieldset>
¿Eres un humano?<br />
<script type="text/javascript">DrawBotBoot()</script>
```



```
<!--<input id="Button1" type="button" value="Check" onclick="ValidBotBoot();" />-->
<p><input type="button" value = "Ingresar" id="btnIngresar"></p>
</fieldset>
</article>
</section>
</body>
<!-- TERMINA LA SEMANTICA DE LA PAGINA -->
<!-- INICIA CODIFICACION EN JAVASCRIPT-->
<script type="text/javascript">
//OCULTA EL ELEMENTO CON EL ID btnIngreso
const btn1 = document.getElementById('btnIngresar').style.visibility="hidden";
</script>
<!-- TERMINA CODIFICACION EN JAVASCRIPT-->
</html>
```

---

## Capítulo 6

### Conclusiones y Trabajos futuros

## 6.1. Conclusiones

Los algoritmos genéticos se encuentran entre las técnicas de aprendizaje automático más recientes, su integración con pruebas de turing completamente automáticas y públicas para diferenciar computadoras de humanos representa una innovación, así como una mayor facilidad para los usuarios que requieren identificarse.

Cuando un sistema web solicita ingresar un código captcha, dependiendo de la forma en que se presente, ya sea visualizar imágenes y relacionarlas u observar letras y números distorsionados que se tienen que descifrar para completar extensos formularios que resultan cada vez más complicados, derivado de que los programadores aplican cada vez más variaciones, distorsiones e introducen más ruido para dificultar la resolución automática por parte de las computadoras.

Desgraciadamente, los captchas que normalmente se emplean presentan problemas de accesibilidad para algunos grupos de usuarios. En la actual investigación, se implementó una prueba de turing con AG, la cual le solicita al usuario formule una palabra con los caracteres definidos, siempre y cuando se cumpla el criterio, el AG llegará a su objetivo y según los resultados arrojados mediante los experimentos de comparación, en menor tiempo de lo que un captcha actual puede resolverse, lo que resulta ser efectivo y práctico para el usuario y a su vez garantiza que quién inicia sesión en el Sistema Integral del Tecnológico de Misantla (SITM) es un humano o persona real y no un algoritmo de spam.

Como producto de esta investigación se logró realizar la integración de un algoritmo evolutivo con una prueba de Turing, enfocado a un ámbito web para el Sistema Integral del Tecnológico de Misantla y así garantizar que es un humano el que trata de ingresar y no un algoritmo automático.

Para resumir, como resultados se han obtenido: La implementación e implantación el subsistema de seguridad para el control de acceso y transferencia de datos del Sistema Integral del Tecnológico de Misantla.

- La integración de los algoritmos genéticos como prueba de Turing completamente automática y pública para diferenciar ordenadores de humanos.
- Una innovación y una mayor facilidad para los usuarios que requieren iden-

tificarse, ya que al ingresar el código captcha en el formato que se presenta, en la actual investigación, no se tuvo la necesidad de intentar identificar el código que se presenta, así como también visualizar imágenes y relacionarlas a imágenes, métodos que son para los usuarios tediosos y complicados.

- Con la única instrucción de ingresar una palabra, cumpliendo una sola instrucción, en este caso escribir una palabra con los caracteres que se le muestra en la página, no requiere tener una mayor concentración, como hacer sumas, cálculos o relaciones, más que escribir una sola palabra.

## **6.2. Trabajos futuros**

Es claro que el trabajo con el avance de las tecnologías de programación pueda evolucionar y es así que, por lo antes mencionado, se tienen como trabajos futuros se proponen los siguientes:

- Generación aleatoria de nomenclatura a validar.
- Reconocimiento facial como capa de seguridad agregada.

---

# Bibliografía

- [1] Pressman S. R., Ingeniería del Software un enfoque práctico, McGrawHill, pp. 33-43, (2010) Fernández, Y., Díaz Y., Patrón Modelo-Vista-Controlador, Revista Telem@tica, 11(1),
- [2] Fernández, Y., Díaz Y., Patrón Modelo-Vista-Controlador, Revista Telem@tica, 11(1),
- [3] «Acunetix Vulnerability Scanner: Web Application Security», Acunetix. 2018.
- [4] «Nmap: the Network Mapper - Free Security Scanner». [En línea]. Disponible en: <https://nmap.org/>. [Accedido: 02-may-2018].
- [5] «¿Qué es un Certificado SSL? — Cert Superior». [En línea]. Disponible en: <https://www.certsuperior.com/QueesunCertificadoSSL.aspx>. [Accedido: 02-may-2018].
- [6] «¡ Que es un CAPCHA? de Maria Betania Labus en Prezi». [En línea]. Disponible en: <https://prezi.com/lu5lyhshqrra/que-es-un-capcha/>. [Accedido: 04-may-2018].
- [7] «¿Qué es CAPTCHA?» [En línea]. Disponible en: <https://www.aboutespanol.com/que-es-captcha-157845>. [Accedido: 04-may-2018].
- [8] MSc. Susana C. Romaniz. (s. f.). Seguridad de aplicaciones web: vulnerabilidades en los controles de acceso. Grupo de Investigación en Seguridad de las Tecnologías de Información y Comunicaciones Facultad Regional Santa Fe - Universidad Tecnológica Nacional.

- [9] Ortega Martorell, S., y Canino Gutiérrez, L. (2006). PROTOCOLO DE SEGURIDAD SSL. Ingeniería Industrial, XXVII(2-3). Recuperado a partir de <http://www.redalyc.org/resumen.oa?id=360433561012>
- [10] Muro, J. D., y Ramírez, E. L. (2000). Seguridad y comercio por internet. Conciencia Tecnológica, (13). Recuperado a partir de <http://www.redalyc.org/articulo.oa?id=94401307>
- [11] Vega, O. A., y Vinasco-Salazar, R. E. (2014). CAPTCHA: ¿Una solución para la seguridad informática o problema para la accesibilidad/usabilidad web? Revista e-Ciencias de la Información, 4(2), 1-14.
- [12] López, J. A. A., y Rodríguez, C. R. R. (2016). Seguridad en internet. Paakat: Revista de Tecnología y Sociedad, (11). Recuperado a partir de <http://www.redalyc.org/articulo.oa?id=499054323006>
- [13] Nodarse, F. A. F. (2013). Sobre Comercio electrónico en la WEB 2.0 y 3.0. Revista Cubana de Ciencias Informáticas, 7(3), 96-113.
- [14] PHP: ¿Qué es PHP? - Manual. (s. f.). Recuperado 20 de febrero de 2018, a partir de <http://php.net/manual/es/intro-what-is.php>
- [15] HTML 5.2. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.w3.org/TR/html52/>
- [16] JSON. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.json.org/json-es.html>
- [17] JavaScript. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://developer.mozilla.org/es/docs/Learn/JavaScript>
- [18] Sublime Text 3 Beta ya disponible. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://web.archive.org/web/20141027110117/http://falasco.org/sublime-text-3-beta-ya-disponible>
- [19] About the XAMPP project. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.apachefriends.org/es/about.html>

- [20] MySQL. (s. f.). Recuperado 20 de febrero de 2018, a partir de <http://www.oracle.com/technetwork/database/mysql/index.html>
- [21] JSON. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.json.org/json-es.html>
- [22] HTML 5.2. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.w3.org/TR/html52/>
- [23] ¿Qué es Ubuntu? (s. f.). Recuperado 20 de febrero de 2018, a partir de <http://www.ubuntumx.org/queesubuntu.php>
- [24] 1.2. ¿Qué es GNU/Linux? (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.debian.org/releases/stable/armel/ch01s02.html.es>
- [25] ¡Bienvenido! - El proyecto del servidor Apache HTTP. (s. f.). Recuperado 20 de febrero de 2018, a partir de <http://httpd.apache.org/>
- [26] Hacking (Hacker): qué es y cómo proteger su PC — Avast. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.avast.com/es-es/c-hacker>
- [27] Standard ECMA-262. (s. f.). Recuperado 20 de febrero de 2018, a partir de <http://www.ecma-international.org/publications/standards/Ecma-262.htm>
- [28] Modelo de objetos de documento W3C. (s. f.). Recuperado 20 de febrero de 2018, a partir de <https://www.w3.org/DOM/>
- [29] Just Sherekan - Blog de Programación, Introducción a Ajax. (2008, abril 26). Recuperado 20 de febrero de 2018, a partir de <https://web.archive.org/web/20080426031910/http://sherekan.com.ar/blog/2008/04/19/introduccion-a-ajax/>
- [30] «Revista Derecho penal y criminología UExternado». [En línea]. Disponible en: <https://revistas.uexternado.edu.co/index.php/derpen/article/view/965>. [Accedido: 16-oct-2018].

- 
- [31] Caballero, J. J. V. (2016). Modelo de procesos para el desarrollo del front-end de aplicaciones web. <http://alicia.concytec.gob.pe/vufind/Record/RULIc3d0a75d403c26d921004c6248c07423>. INTERFASES.
- [32] de Guevara, L. V. (2018). Gestión de bases de datos. <https://media.readthedocs.org/pdf/gestionbasesdatos/latest/gestionbasesdatos.pdf>.
- [33] Marcotte, E. (2011). Responsive web design (2da Edition ed.). A Book Apart.
- [34] Bahit, E. (2011). El paradigma de la programación orientada a objetos en php con el patrón arquitectónico mvc.
- [35] Castejón, J. S. (2004). Arquitectura y diseño de sistemas web modernos. *INFORMAS* , 1 , 2-3.
- [36] Pressman, R. S. (2010). Ingeniería del software (7ma Edition ed.). McGraw-Hill.
- [37] Boehm, B. W. (1988). A spiral model of software development and enhancement. *IEEE Computer Society* , 21 , 61 - 72.
- [38] Fernández, D. Y., Y. (2012). Patrón modelo-vista-controlador. *Revista Telemática*.
- [39] Modelo Espiral de un proyecto de desarrollo de software. (s. f.), 9.
- [40] A. Brindle (1991). Genetic algorithms for function optimization. Tesis doctoral, Universidad de Alberta, Canada.