



**INSTITUTO POLITÉCNICO NACIONAL**  
**ESCUELA SUPERIOR DE CÓMPUTO**  
**ESCOM**

**TRABAJO TERMINAL**

**“Sistema de Búsquedas Personalizadas por Medio de Estructuras de Agentes SURFACE”**

*Que para cumplir con la opción de titulación curricular en la carrera de*

**“Ingeniería en Sistemas Computacionales con especialidad en Sistemas”**

*Presentan*

**Garrido Gutiérrez Miguel Ángel**  
**Pizarro Hernández Sergio Israel**  
**Quintana Mendivil Rodolfo Ivan**  
**Simón Trejo Nancy**

*Directores*

**Dra. Fabiola Ocampo Botello**  
**Ing. José Sánchez Juárez**



**México D.F., 3 de Mayo de 2012**



# INSTITUTO POLITÉCNICO NACIONAL

## ESCUELA SUPERIOR DE CÓMPUTO



No de Registro TT 20110006

Abril 2012

### Documento Técnico

## “Sistema de Búsquedas Personalizadas por Medio de Estructuras de Agentes SURFACE”

### *Presentan*

Garrido Gutiérrez Miguel Ángel<sub>1</sub>  
Pizarro Hernández Sergio Israel<sub>2</sub>  
Quintana Mendivil Rodolfo Ivan<sub>3</sub>  
Simón Trejo Nancy<sub>4</sub>

### *Directores*

Dra. Fabiola Ocampo Botello  
Ing. José Sánchez Juárez

**Advertencia:** Este documento contiene información desarrollada por Instituto Politécnico Nacional en la Escuela Superior de Computo a partir de datos y documentos con derecho de propiedad y por tanto su uso queda restringido a las aplicaciones que explícitamente se convengan.

### **Resumen**

**Sistema de Búsquedas Personalizadas por Medio de Estructuras de Agentes SURFACE** será un sistema que aplicará estructuras de agentes y recursos de metadatos, con el objetivo de permitir a los usuarios realizar búsquedas de información con base en los parámetros proporcionados por el mismo, tales como: palabras clave, direcciones URL, nombres de artículos específicos y/o temas en concreto; para fines del sistema, dichos elementos mencionados, conformarán lo que se define como “Búsquedas Personalizadas”.

Conforme al uso que se le vaya dando al sistema, este será capaz de almacenar el historial de paginas buscadas utilizadas con una antigüedad de dos semanas, así como las fuentes visitadas; dichos criterios son los que permitirán afinar los resultados de la búsqueda, con la intención de disminuir el número de resultados obtenidos y eliminar duplicidad de fuentes.

1 [ronaldomagg@msn.com](mailto:ronaldomagg@msn.com)

2 [sergiopizarrohdz@hotmail.com](mailto:sergiopizarrohdz@hotmail.com)

3 [cubicuate\\_ivan@hotmail.com](mailto:cubicuate_ivan@hotmail.com)

4 [hola\\_nancy1cm5@hotmail.com](mailto:hola_nancy1cm5@hotmail.com)



ESCUELA SUPERIOR DE CÓMPUTO  
SUBDIRECCIÓN ACADÉMICA  
DEPARTAMENTO DE FORMACIÓN INTEGRAL E INSTITUCIONAL



COMISIÓN ACADÉMICA DE TRABAJO TERMINAL

México D.F. a 3 de Mayo del 2012.

**ING. APOLINAR FCO. CRUZ LÁZARO**  
**PRESIDENTE DE LA COMISIÓN ACADÉMICA DE**  
**TRABAJO TERMINAL**  
**PRESENTE**

Por medio del presente, le informo que los alumnos que integran el TRABAJO TERMINAL 2001-0006 titulado **“Sistema de Búsquedas Personalizadas por medio de Estructuras de Agentes SURFACE”** concluyeron satisfactoriamente su trabajo.

El empastado del reporte Técnico Final y el Disco Compacto (CD) fueron revisados ampliamente por un servidor y corregidos, cubriendo el alcance y el objetivo planeados en el protocolo original y de acuerdo a los requisitos establecidos por la Comisión que usted preside.

**ATENTEMANTE**

---

**Dra. Fabiola Ocampo Botello**

Director

---

**Ing. José Sánchez Juárez**

Director

## ADVERTENCIA

*“Este informe contiene información desarrollada por la Escuela Superior de Cómputo del Instituto Politécnico Nacional a partir de datos y documentos con derecho de propiedad y por lo tanto su uso quedar restringido a las aplicaciones que explícitamente se convengan.”*

La aplicación no convenida exime a la escuela su responsabilidad técnica y da lugar a las consecuencias legales que para tal efecto se determinen.

Información adicional sobre este reporte técnico podrá obtenerse en:

En La Subdirección Académica de la Escuela Superior de Cómputo del Instituto Politécnico Nacional, situada en Av. Juan de Dios Bátiz s/n  
Teléfono: 57296000 Extensión 52000

## **Agradecimientos.**

A nuestros padres principalmente por la ayuda incondicional en los momentos más difíciles de la carrera, a todas y cada una de las personas que sin esperar nada a cambio compartieron pláticas, conocimientos y diversión.

A nuestros directores: Dra. Fabiola Ocampo Botello e Ing. José Sánchez Juárez que nos brindaron su tiempo y conocimiento para realizar este trabajo terminal  
A todos los que ayudaron a convertir este sueño en realidad.  
¡MUCHAS GRACIAS!

# Índice

Capítulo 1 Introducción.....	9
Capítulo 2 Antecedentes.....	11
2.1 Introducción .....	11
2.2 Estado del arte .....	12
2.2.1 Software disponible en el mercado .....	12
2.3 Marco histórico .....	13
2.4 Marco teórico .....	14
2.4.1 Agentes.....	14
2.4.3 Buscador Web .....	17
2.4.4 Servidores.....	17
2.4.5 Alojamiento Web .....	18
4.5 Motores de Búsqueda.....	18
Capítulo 3 Problemática.....	20
3.1 Planteamiento del problema .....	20
3.2 Propuesta de solución.....	20
3.3 Objetivos .....	21
3.3.1 Objetivo General .....	21
3.3.2 Objetivos Específicos .....	21
Capítulo 4 Análisis del sistema .....	22
4.1 Estudio de Factibilidad.....	22
4.1.1 Factibilidad Operativa .....	22
4.1.2 Factibilidad Técnica .....	22
4.1.3 Estudio de Factibilidad Económica.....	23
4.1.4 Estudio de Factibilidad Legal.....	23
4.3 Análisis de Riesgos .....	23
4.4 Estudio de software .....	26
4.5 Arquitectura de Software .....	27
4.6 Modelo de Arquitectura .....	27
4.7 Estilo de Arquitectura.....	28
4.8 Metodología de Desarrollo Web .....	29
Capítulo 5 Diseño del sistema.....	31
5.1 Requerimientos Funcionales y No Funcionales .....	31

5.2 Diagrama a bloques .....	32
5.3 Definición de casos de uso .....	33
5.4 Diagrama de Búsqueda y Filtrado .....	43
5.6 Diagrama de Actividades .....	60
5.7 Diagrama de Estado .....	61
5.8 Diagrama Relacional .....	62
5.9 Pantallas .....	64
Referencias .....	66
Citas.....	66
Antecedentes .....	66
Fuentes de marco histórico.....	66
Análisis del sistema.....	67
Fuentes de arquitectura de Software .....	67
Fuentes de modelo de arquitectura que se usara .....	67
Fuentes de estilo de arquitectura .....	67
Lecturas Recomendadas.....	67
Referencias de Autores.....	68

## Índice de Figuras

Figura 1: Grafica de precisión de búsqueda comparativa Google/Bing[1] .....	11
Figura 2: Grafico comparativo de uso de buscadores web [4].....	17
Figura 3: Diagrama a Bloques.....	32
Figura 4: Diagrama de Casos de Uso .....	33
Figura 5: Diagrama de Búsqueda y Filtrado .....	43
Figura 6: Diagrama de Secuencia Crear Perfil .....	44
Figura 7: Diagrama de Secuencia: Iniciar Sesión.....	45
Figura 8: Diagrama de Secuencia: Modificar imagen de perfil.....	46
Figura 9: Diagrama de Secuencia: Modificar datos de cuenta.....	47
Figura 10: Diagrama de Secuencia: Modificar datos personales .....	48
Figura 11: Diagrama de Secuencia: Crear perfil personalizado de búsqueda.....	49
Figura 12: Diagrama de Secuencia: Modificar perfil personalizado de búsqueda.....	50
Figura 13: Diagrama de Secuencia: Eliminar perfil personalizado de búsqueda .....	51
Figura 14: Diagrama de Secuencia: Registrar temas.....	52
Figura 15: Diagrama de Secuencia: Eliminación temas.....	53
Figura 16: Diagrama de Secuencia: Registrar URL .....	54
Figura 17: Diagrama de Secuencia: Eliminación URL .....	55
Figura 18: Diagrama de Secuencia: Registro de Favoritos .....	56
Figura 19: Diagrama de Secuencia: Eliminación de Favoritos.....	57

Figura 20: Diagrama de Secuencia: Realizar búsqueda con registro .....	58
Figura 21: Diagrama de Secuencia: Realizar búsqueda sin registro.....	59
Figura 22: Diagrama de Actividades.....	60
Figura 23: Diagrama de Estados.....	61
Figura 24: Diagrama Relacional.....	63
Figura 25: Pantalla de Búsqueda.....	64
Figura 26: Pantalla de Búsqueda Usuario Registrado .....	64
Figura 27: Pantalla de Registro de Perfil .....	65
Figura 28: Pantalla de Registro de Usuario .....	65
Figura 29: Pantalla de Resultados de Búsqueda .....	65

## Índice de Tablas

Tabla 1: Tabla comparativa de buscadores.....	12
Tabla 2 Relación Problema/Solución.....	22
Tabla 3: Comparación entre servidores .....	22
Tabla 4: Riesgos.....	24
Tabla 5: Contingencia.....	25
Tabla 6: Tabla comparativa de Software.....	26

# Capítulo 1 Introducción

En octubre de 1962, DARPA instaló una red de computadoras en distintas partes de los Estados Unidos, esta red, conformada por cuatro computadoras, comunicaba a los usuarios entre ellos con el centro de operaciones y servía para la investigación e intercambio de información sobre “ordenadores avanzados”, el problema de no tener una red común entre las máquinas se hizo evidente, puesto que desde la máquina principal solo se podía comunicar con 1 de las otras 3 y si se necesitaba comunicarse con otra de las 2 se necesitaba desconectar e iniciar con una sesión nueva.

En 1966, la agencia de proyectos de investigación avanzada (Advanced Research Projects Agency, ARPA) tenía un programa con varias instituciones de investigación. El objetivo de ARPA era enlazar diferentes ordenadores todos juntos, para mejorar la potencia general de los procesamientos de los ordenadores y descentralizar el almacenamiento de información.

El 21 de noviembre de 1969, se creó el primer enlace entre las universidades de UCLA y Stanford, esto se llevó a cabo por medio de una red telefónica conmutada y posteriormente el mismo año se llevó a cabo en la Universidad de Michigan una red basada en la conmutación de paquetes con un protocolo de nombre “X.25”.

Para finales de 1970 ARPANET empezó con el manejo de un protocolo que ahora conocemos como TCP/IP, protocolo que impulsó el crecimiento explosivo en el uso de ARPANET y el desarrollo de muchas herramientas y utilidades que ahora podemos conseguir gratuitamente.

Cuando este protocolo se empezó a implementar de manera más estable, en 1984, se convenció al CERN de utilizar este protocolo para crear una red. El CERN permaneció desconectado del resto de las redes formando una pequeña red interna. Luego de esto en 1988 mientras se mantenían pláticas sobre la conversión de la UUCP, red europea, a TCP/IP el CERN abre su primera red externa al resto del mundo por medio de los routers de la entonces pequeña compañía CISCO.

Desde entonces el Internet pasó de ser una herramienta de algunos selectos lugares de investigación científica, para el intercambio interno de información, a una herramienta que se usa alrededor del mundo para búsqueda de información en línea.

En las últimas tres décadas, el auge del Internet como fuente de búsqueda de información ha tenido un crecimiento exponencial, debido a su gran repositorio de información y la gran cantidad de fuentes que éste nos ofrece. Sin embargo, esta gran cantidad de datos, la cual aumenta a cada momento, ha hecho que el encontrar información cuya fuente sea de confianza para el usuario, se dificulte mucho más.

SURFACE ofrecerá búsquedas las cuales utilizarán filtrados, que reducirán la cantidad de datos que se mostraran al usuario, mediante el uso de estructuras de agentes, el

sistema podrá, no solo reducir el número de fuentes que se muestren, sino también tomar en cuenta las preferencias que el usuario proporcionadas mediante la creación un perfil.

Actualmente el buscador más usado a nivel mundial es GOOGLE, puesto que sus algoritmos de búsqueda han sido mejorados en sus trece años de vida, ofreciendo excelentes resultados, sin embargo la repetición de fuentes se ha convertido en un problema que parece ocurrir en cada búsqueda realizada. Por otro lado, el buscador Copernic, ofrece búsquedas confiables mediante sus estructura de motores de búsqueda, la cual está conformada por 30 en su versión libre y más de 50 en la versión de paga, esto nos da como resultado un poderoso buscador, sin embargo su portabilidad y la cantidad de recursos máquina que ocupa dejan en desventaja a este buscador frente a los buscadores basados en web.

Tomando en cuenta todo lo antes mencionado, SURFACE toma como base las ventajas de otros sistemas, para así ofrecer al usuario una mejor experiencia en la navegación, mediante el uso de los recursos antes mencionados, más la ventaja de que será una aplicación basada en web, permitiéndole al usuario acceder a su perfil desde cualquier parte del mundo, reduciendo la utilización de recursos, sin mencionar que el sistema operativo que utilice el usuario no afectará al sistema.

El perfil de usuario no solo contendrá información sobre el usuario, como su correo electrónico, nacionalidad, etc. si no que también tendrá la opción para que este agregue sus propias fuentes que considere de confianza, esto ayudará a los filtros de información y hará de cada búsqueda una experiencia completamente personalizada.

## Capítulo 2 Antecedentes

### 2.1 Introducción

En los últimos años se han creado numerosos buscadores, que nos ayudan a encontrar información en la red, estos buscadores implementan diversos algoritmos matemáticos y estructuras de agentes y en algunos casos motores de búsqueda preexistentes, sin embargo, la calidad de los resultados no depende del algoritmo utilizado o de la cantidad de resultados, sino de la calidad de los contenidos.

El buscador más usado a nivel mundial es GOOGLE por su rapidez, su renombre, entre otros factores, sin embargo estudios recientes demostraron que el buscador Bing, de Microsoft, ha superado en la precisión, el acceso a las páginas después de realizar búsquedas, es un indicativo de la precisión de los buscadores.

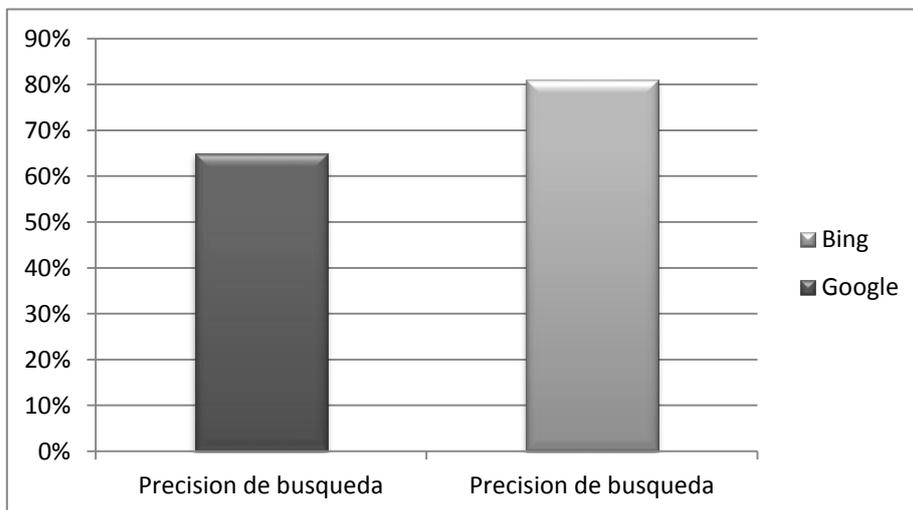


Figura 1: Grafica de precisión de búsqueda comparativa Google/Bing[1]

Un estudio realizado por la compañía Experian Hitwise ha monitorizado las búsquedas de 10 millones de internautas en Estados Unidos y quedó demostrado, como se puede ver en la gráfica, que en el 81% de las veces que se realizó una búsqueda en BING se accedió a una página, mientras que en GOOGLE solo se accedió a un 65% de las páginas que se desplegaron. Sin embargo la mayor parte de las búsquedas se realizó por GOOGLE.

Debido a la gran cantidad de información que se encuentra en Internet, se desarrollará un buscador que implemente búsquedas personalizadas, de este modo SURFACE implementará perfiles de navegación, agentes inteligentes y algoritmos de filtrado, para disminuir el número de resultados obtenidos a partir de una exploración en la red.

Mediante la utilización del Perfil de navegación, previamente creado por el usuario, se realizará un filtrado de la información obtenida a partir de una búsqueda, el filtrado se

realizará empleando motores de búsqueda. Así mismo, se utilizarán estructuras de agentes para la implementación de filtros, dichos filtros se basarán en el uso que le vaya dando el usuario al sistema, apegándose, a las preferencias de búsqueda del usuario.

## 2.2 Estado del arte

### 2.2.1 Software disponible en el mercado

En la actualidad en el mercado se pueden encontrar distintos tipos de buscadores especializados en información científica y técnica, sin embargo no hay alguno que sea de uso común y que se guíe por las preferencias del usuario permitiendo hacer las búsquedas más personalizadas, en la Tabla 2.2 enlistamos algunos de los productos en el mercado, así como sus características principales y su precio en el mercado.

Tabla 1: Tabla comparativa de buscadores

Software Comercial	Descripción	Características
<b>SCIRUS</b>	Buscador en la Web especializado en información científica –técnica	Recupera páginas web, además artículos de revistas. Para el acceso a los artículos es necesario autenticarse. Efectúa búsqueda avanzada que contiene más opciones de búsqueda para mayor precisión. Se puede seleccionar la fuente de información donde se realizara la búsqueda, se permite seleccionar el tipo de información que se desea y permite la limitación por años de publicación
<b>GOOGLE Scholar</b>	Herramienta que indexa documentación técnica y ofrece resultados de búsqueda dirigidos a la comunidad científica y académica	Los documentos que muestra están ordenados en función de la importancia de la publicación en la que se encuentra el artículo o el autor o frecuencia en la que es citado el escrito. Dispone de un menú de opciones avanzadas en las que se puede especificar el autor de los artículos, la publicación o año de publicación.
<b>Copernic Agent Personal</b>	Ofrece la obtención de resultados relevantes agrupados en categorías.	Da mejores resultados del motor de búsqueda mediante la consultas de múltiples motores de búsqueda a la vez, la combinación de sus resultados, la eliminación de duplicados y mantener solo lo mejor de la información obtenida de los motores de búsqueda a consultar
<b>Citeseer</b>	Es una biblioteca digital de literatura científica	Mejora la difusión y la retroalimentación de la literatura científica. Además de la indexación del artículo principal, las citas incluidas en artículos también son indexadas.
<b>SURFACE</b>	El sistema de búsquedas personalizadas, mediante el uso de perfiles de usuario, estará basado en el uso de estructuras de agentes.	Usará tecnología de agentes y por medio de la red neuronal artificial encontrará resultados más personalizados a lo que el usuario busque en la web. Esto por medio de un perfil de usuario, historial de visitas del usuario y por sus criterios de búsqueda que pueden ser artículos, URL y palabras claves.

## 2.3 Marco histórico

El primer buscador de Internet ampliamente usado fue NSCA Mosaic siendo este uno de los primeros buscadores web gráficos de Internet ya que proporciona un acceso muy sencillo a Internet. En 1994 la revista Fortune distinguió a Mosaic nombrándolo el producto del año 1994. Describiendo al buscador como un programa que transformaba a Internet en una red manejable que conecta con sencillez a usuarios y servidores. Mosaic fue desarrollado en el centro nacional para aplicaciones de computación (*National Center for Supercomputing Applications, NCSA*). Fue diseñado por estudiantes y se puso a disposición del público en forma gratuita.

Opera nació como el proyecto independiente de dos programadores de la compañía noruega Telenor. Jon Tetzschner y Geir Ivarsoey invirtieron 7000 dólares de su propio bolsillo y seis meses de trabajo continuo de un prototipo con la intención de crear el buscador más rápido de la tierra, la primera versión fue liberada al mundo en 1995. Opera se desmarcó del resto por su limpieza y su capacidad de innovación, introduciendo nuevos conceptos como las pestañas o 'tabs'. Hasta ahora, la versión libre de publicidad costaba 39 dólares al año. Se calcula que tiene unos 22 millones de usuarios en el mundo. Desde que se convirtió en gratuito, ha recibido cinco millones de descargas.

Más tarde apareció el buscador Netscape Navigator fue diseñado por Mac Andreessen, que creó el prototipo de NCSA Mosaic siendo estudiante en NCSA. Una vez graduado, él y unos compañeros, fueron contratados por Jim Clark, ex director general de Silicon Graphics, para crear una nueva empresa de informática cuyo principal objetivo sería desarrollar programas para la Web. Netscape Navigator incluye un marcador que indica el porcentaje del proceso de transferencia que ya ha sido realizado. El programa indica cuando consigue conectarse a un servidor, visualizar el número total de bytes que hay que importar y mantener la cuenta de los que ya han sido recibidos. Al recibir imágenes de las páginas iniciales lo hace por partes, esto es, en lugar de esperar a que aparezca la imagen completa, comienza por visualizar una imagen de baja resolución y, posteriormente, la actualiza tres o cuatro veces hasta que se ve perfecta. Netscape maneja la lista de *bookmarks* ya que se visualizan como opciones adicionales de uno de los menús. Netscape Navigator también es compatible con ciertas extensiones del lenguaje HTML utilizado para la elaboración de páginas Web que, en la actualidad, emplean ya muchos servidores de Web.

Pero la idea no era hacer una simple copia, sino algo mejor, que corrigiera los errores y meterle todo lo que no se había podido, iba a ser un Mosaic Killer que pronto se abreviaría Mozilla.

En Octubre de 1994, Netscape publico la primera versión beta de su buscador en Internet, el Mozilla 0.96b. En diciembre, la versión final ya estaba preparada, Mozilla 1.0, convirtiéndose en el primer buscador Web comercial. Una vez el código abierto del

buscador Netscape fue desplegado en 2002, también llamado Mozilla y posteriormente publicada como Firefox en noviembre de 2004.

Netscape fue un éxito por lo que Microsoft compro los derechos de Mosaic, lo mejoro y saco su propio buscador, Internet Explore para el sistema operativo Microsoft Windows desde 1995. Este buscador fue uno de los más utilizados de Internet desde 1999 hasta la actualidad.

Konqueror es un buscador libre y parte oficial del proyecto KDE. Funciona como gestor de archivos, buscador web o visor de archivos. Permite la navegación en directorios locales, pre visualiza cualquier documento o archivo, incluyendo imágenes y videos, según los expertos el mejor buscador de la historia. Especialmente diseñado para Linux.

Safari es el estándar de navegación de Mac OS X, el sistema operativo de Apple. Se presentó en el Macworld del 2003 y supuso la ruptura total con Microsoft, quien todavía desarrollaba Internet Explorer para Apple. Implementa un lector de fuentes RSS y puede guardar la página completa en un archivo para verla sin que el usuario esté conectado a la Red.

Por último y el más utilizado hasta la actualidad el buscador más usado es Google Chrome de código abierto desarrollado por Google. Este buscador está constituido por componentes de buscadores de código abierto como son WebKit y Mozilla. El objetivo del buscador es lograr una interfaz sencilla, estabilidad, velocidad y seguridad. La versión beta para Windows fue lanzada el 2 de Septiembre del 2008, próximamente lanzaran las versiones para Linux y Mac OS X.

## **2.4 Marco teórico**

### **2.4.1 Agentes**

Actualmente en el área de la Inteligencia Artificial ha surgido un nuevo paradigma conocido como “paradigma de agentes”. Este paradigma aborda el desarrollo de entidades que puedan actuar de forma autónoma y razonada.

Los agentes de información o agentes de Internet son sistemas de software que tienen acceso a múltiples fuentes de información en el mundo [Klusck 1999]. Esta clase de agentes ayudan a los usuarios en la búsqueda de información útil y relevantes, además de proveer acceso a diferentes fuentes de información debe de ser capaz de recuperar, analizar, manipular, e integrar la información cuando es solicitada.

De forma específica se puede destacar los siguientes tipos agentes:

- **Agentes Autónomos.-** Se trata de un programa que viaja entre los sitios Web, decidiendo por el mismo que debe hacer y cuando debe moverse a otros lugares.
- **Agentes Inteligentes.-** Se trata de un programa que ayuda al usuario a ciertas acciones.

- **Agentes de Usuario.-** El nombre técnico para denominar a un programa que ejecuta determinadas tareas para un usuario en la red.
- **Agentes de Información.-** Este tipo de agentes pueden asistir al usuario en la búsqueda y filtrado de información relevante, informar cuando nueva información relevante está disponible. Por lo que en cada uno de los casos los agentes de información ayudan al usuario en la ejecución de tareas estas se pueden llevar a cabo de manera independiente o trabajar en coordinación con otros agentes, por lo que tienen que ser capaces de capturar y almacenar las preferencias del usuario.

Hay diferentes maneras de clasificar a los agentes de información.

La primera clasificación toma como criterio el ámbito de actuación de los agentes [Caglayan, 1997], distinguiendo entre:

- **Agentes de escritorio:** Se encarga de tareas habituales en el manejo del ordenador. Depende del tipo de software que sirva como base, se puede distinguir entre agentes de sistema operativo, agente de aplicaciones.
- **Agentes Intranet:** Operan en una red de área local y facilitan el uso de la red, el compartir información y la cooperación entre los usuarios. Entre ellos se encuentran los agentes de ayuda colaborativa, agentes de bases de datos, agentes de automatización de procesos.
- **Agentes Internet:** Su ámbito de trabajo es Internet. En función del tipo de acciones que realiza, se puede encontrar agentes de búsqueda, filtrado, recuperación de información, agentes de notificación, agentes móviles.

En [Klusck 1999] se propone cinco clases de agentes de información, atendiendo a las características que poseen como agentes:

- **Agentes cooperativo.-** Cuando los problemas son complejos los agentes pueden cooperar entre sí para realizar una parte del problema.
- **Agentes no cooperativo.-** Son agentes individuales que dan un servicio a un usuario concreto cuyos hábitos son aprendidos por el agente como el objetivo de ajustarse a sus necesidades de Información.
- **Agentes adaptivos.-** Son capaces de adaptarse por sí mismos a cambios en su entorno. Los aspectos importantes de este tipo de agentes son que en Internet los agentes tienen que construirse para tratar de una manera fiable y segura con información incierta e incompleta. Otro es la personalización, ya que el sistema se adapta a cada usuario.

- **Agentes racionales.-** Son capaces de actuar por si mismos o de colaborar con el fin de incrementar sus capacidades de búsqueda.
- **Agentes móviles.-** Son capaces de navegar de forma autónoma a través de Internet de un sitio a otro para la ejecución de sus tareas en diferentes servidores.

La última clasificación en la que se pueden dividir los agentes de información es según su función.

**Agentes de búsqueda:** Son agentes que buscan, recuperan y proporcionan la información como si fueran auténticos gestores de información y documentación.

Estos agentes nos proporcionan una facilidad de uso, y un incremento de la productividad si el usuario sabe lo que desea, un incremento de la precisión en la búsqueda sin aumentar notablemente el número de documentos devueltos y una reducción de la sobrecarga que genera los procesos de búsqueda en la web.

**Agentes de filtrado:** Son aquellos que se usan para reducir la sobreabundancia de información mediante el borrado de los datos no deseados. Este tipo de agentes necesita almacenar, aprender y manipular las preferencias y gustos de cada uno de ellos.

Los agentes de filtrado consisten en determinar si un artículo es relevante o irrelevante para el usuario basándose en el perfil de usuario [Caglayan, 1997], en un entorno donde la probabilidad de encontrar un documento relevante es muy baja en comparación con la alta probabilidad de encontrar un documento irrelevante.

El funcionamiento de un agente de filtrado de información es el siguiente:

- ✓ Establece el perfil de usuario.  
Esto puede llevarse a cabo de dos formas diferentes:  
**Directa:** Cuando el usuario determina su perfil a través de una interfaz en el que especifica, por medio de palabra claves y temas de interés.  
**Indirecta:** Por el control efectuado por el propio agente de los enlaces seguidos por el usuario en una sesión concreta.
- ✓ Comienza la búsqueda en las diferentes fuentes de información
- ✓ Una vez filtrados aquellos artículos que se ajustan a la especificación del usuario, elabora el informe que presenta los resultados obtenidos los cuales serán enviados a la dirección correspondientes

**Agentes de monitorización:** Proporcionan al usuario la información cuando sucede un acontecimiento determinado. También se pueden considerar como un software localizado en un determinado servidor que se encarga de descubrir y notificar eventos interesantes especificados previamente por el usuario.

### 2.4.3 Buscador Web

Un Buscador Web o Navegador Web, es una aplicación que funciona a través de Internet, que puede interpretar la información que se encuentra en los sitios web para que el usuario pueda leerla.

Un buscador interpreta el código, por lo general escrito en HTML, en que se encuentre escrita la página y le da un interpretación textual o gráfica, permitiendo así al usuario poder interactuar con el contenido de la página o visitar otras páginas mediante hipervínculos o enlaces.

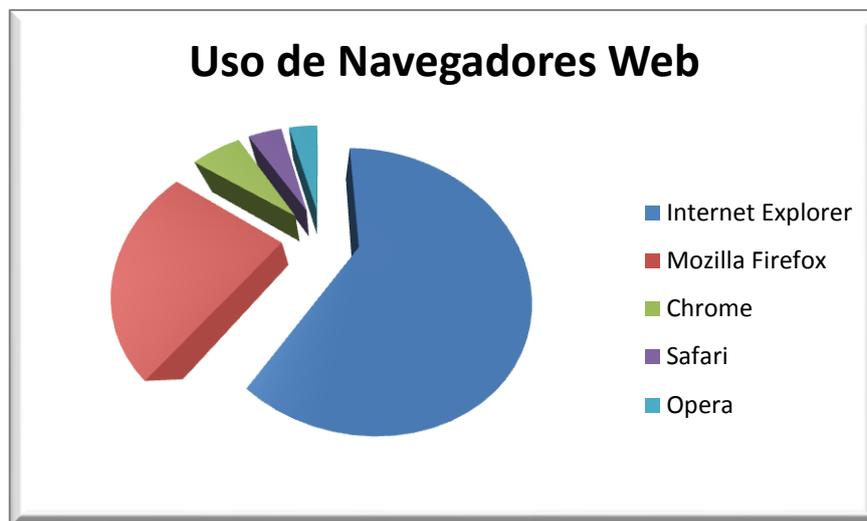


Figura 2: Grafico comparativo de uso de buscadores web [4]

### 2.4.4 Servidores

Cuando se habla de servidores web, nos podemos referir al sistema físico o a la paquetería software encargada de las funciones que controlan las tareas que el servidor físico realiza. Para propósitos del mismo Trabajo Terminal, nos enfocaremos al servidor web, como un ente físico sin embargo no dejaremos fuera la parte software, a la cual haremos referencia más adelante.

El termino servidor se refiere a un tipo de software destinado a realizar tareas que el usuario le destine, sin embargo, también podemos definir al servidor como el ordenador donde este software funciona, este ordenador tiene como propósito principal el proveer de datos a otras máquinas para que estas puedan hacer uso de los datos que esta les entrega.

Los factores más relevantes para la adquisición de un servicio de servidor o de un servidor como tal son los siguientes:

- Procesador
- Memoria RAM
- Modelo de disco duro

Los servidores web son programas informáticos cuya tarea principal es la de alojar sitios y/o aplicaciones, las cuales son accedidas por los clientes mediante el buscador que se comunica con este mediante protocolos HTTP.

Un servidor web se puede definir como un programa que está dedicado a escuchar las peticiones de los usuarios y a atender o satisfacer las mismas. Por medio de la especificación de la solicitud del usuario, un servidor puede ejecutar un programa o buscar una página web en específico, pero necesariamente debe regresar un resultado como producto de la búsqueda.

#### **2.4.5 Alojamiento Web**

El alojamiento web (web hosting, en inglés) es el servicio que se da a los usuarios, mediante un sistema, para el almacenamiento de información, ya sean videos, archivos de texto, imágenes, etc. El alojamiento web nos habla del espacio físico que ocupara una página web, sistema, dirección de correo electrónico, archivo, etc. durante un tiempo determinado.

Los anfitriones o host son ordenadores que tienen la función de ser el punto de inicio y final de las transferencias de datos que se realizan. Este anfitrión tendrá una dirección de Internet (dirección IP) y un nombre de dominio únicos para este.

Por lo general el servicio de hospedaje para una sola persona puede llegar a resultar costoso, por ende este servicio se puede dar, desde una misma localidad, a varios usuarios, para así poder dividir los costos y poder compartir la rápida conexión a Internet a sus diferentes archivos web.

#### **4.5 Motores de Búsqueda**

Los motores de búsqueda siendo una de las herramientas más utilizadas en Internet para la búsqueda de información desde mediados de 1990, son sistemas informáticos que buscan archivos almacenados en servidores, con ayuda de spiders que son programas diseñados para explorar páginas Web en forma automática.

Los motores de búsqueda comparten elementos en común como son:

- Enlaces al sitio Web que contiene la letra buscada (tags).
- Descripción del sitio.
- La dirección URL completa sobre del sitio Web presentado.

Para encontrar la información requerida por el usuario los motores responden a una consulta que se hace en el formulario de búsqueda y estos regresan un listado de sitios que contienen la palabra solicitada anteriormente.

Estas palabras son sometidas a distintos algoritmos según la importancia que los agentes les asignen. A esta parte se le define como ranking y es el más importante para el agente de filtrado.

Existen tres tipos de buscadores

- ✓ **Buscadores.-** La mayoría de los grandes buscadores internacionales utilizan esta tecnología pero se requieren muchos recursos para su funcionamiento.

Los buscadores recorren las páginas recopilando información sobre el contenido de las páginas para ingresarlas a una base de datos.

Cuando se realiza una búsqueda estos consultan su base de datos, la cual se genera por las altas de los usuarios que las crearon, con la información de relevancia que se contiene de la página Web.

Cada cierto tiempo, los motores revisan la red, para actualizar sus contenidos.

Los buscadores tienen una colección de programas simples y potentes con diferentes objetivos. Se pueden dividir en los programas que exploran la red llamados Spiders, y son estos los encargados de poblar la base de datos.

- ✓ **Directorios.-** Es una tecnología barata, debido por la cantidad de programas scripts en el mercado. No se requiere de muchos recursos de informática pero el soporte y mantenimiento que se le da a este tipo de motores es humano.

La lógica para catalogar la información se realiza a través de una organización clara y subdividida en argumentos. No recorren las Web ni almacenan sus contenidos. El registro de los sitios es de algunos datos como son el título y la descripción de la Web que se introducen cuando se registra un sitio.

Los resultados que muestran los directorios son según la información que ingresa estos ya que estas son revisadas por operadores humanos y clasificadas según categorías para que la búsqueda sea hecha en base a las clasificaciones que se les hayan dado.

- ✓ **Sistemas Mixtos.-** Este tipo de motores de búsqueda son una mezcla entre buscadores y directorios. Además de tener características de los buscadores, la información que entrega al usuario son webs registradas en catálogos sobre contenidos que a su vez son divididos en subsecciones como en el caso de los directorios.

## **Capítulo 3 Problemática**

### **3.1 Planteamiento del problema**

En la última década, la información contenida en Internet ha aumentado de forma prácticamente exponencial, por lo cual, el poder encontrar o localizar información afín a las necesidades de los usuarios, se ha convertido en un problema creciente.

La mayoría de las herramientas para la búsqueda de información, que existen actualmente, se enfocan principalmente en la cantidad de resultados que se pueden obtener dejando al usuario la tarea separar la información, ya sea que esta se repita o que provenga de fuentes poco confiables. Debido a esto el encontrar información, muchas veces es demasiado lento o infructuoso, por otro lado, la gran mayoría de buscadores arrojan la información conforme esta fue encontrada, dejando de lado el nivel de coincidencia que tuvo esta con las palabras que el usuario introdujo, para así, poder tomar una decisión que pueda acercarse más a lo que este requiere.

Por otro lado la especificación de las búsquedas que se realizan, limitan la participación del usuario a las palabras clave. En este caso los sistemas de búsqueda son totalmente autónomos en sus procesos, pues no cuentan con opciones para que el usuario pueda especificar más sus necesidades.

### **3.2 Propuesta de solución**

Tomando en cuenta todo lo antes mencionado, se pretende desarrollar un buscador que no solo sea capaz de realizar una búsqueda, si no de ofrecer al usuario opciones que ayuden al buscador a reducir el número de resultados que este regrese.

SURFACE, será un sistema basado en web, el cual podrá, no solo realizara búsquedas en base a las palabras clave de un usuario y devolver resultados relacionados a estas, si no también contara con un registro, el cual, por medio de un formulario, que se pedirá al usuario que conteste, para completar su registro, podrá realizar búsquedas que sean más a fines con las preferencias del mismo.

Por medio del uso de los metadatos que contienen los sitios web, en conjunción con los datos proporcionados previamente en el registro, el sistema devolverá al usuario los resultados que le proporcionen los procesos de filtrado, ordenándolos dependiendo en el índice de relación entre las palabras clave introducidas en la búsqueda, los datos del formulario y los metadatos.

### **3.3 Objetivos**

#### **3.3.1 Objetivo General**

Desarrollar un sistema de búsqueda, aplicando estructuras de agentes para realizar un filtrado de la información en Internet, con base a los datos almacenados, como lo son, sitios visitados y los perfiles de búsqueda personalizados del usuario. Dicho sistema realizara búsquedas mediante palabras claves, temas, nombre de artículos específicos y/o direcciones URL.

#### **3.3.2 Objetivos Específicos**

- Implementación de búsquedas personalizadas.
- Desarrollar una interfaz grafica mediante la cual se realicen búsquedas en Internet utilizando estructuras de agentes.
- Desarrollar estructuras de agentes las cuales sean capaces de realizar filtrados de la información en Internet.
- Diseñar y desarrollar una base de datos en la cual se almacenen los perfiles de búsqueda de los usuarios.
- Diseñar una estructura de agentes capaz de organizar los datos resultantes de cada ingreso al sistema, con los criterios de búsqueda del usuario.

## Capítulo 4 Análisis del sistema

### 4.1 Estudio de Factibilidad

#### 4.1.1 Factibilidad Operativa

La factibilidad operativa permitirá decidir si el sistema se pondrá en marcha, basándonos en las necesidades que necesitan ser cubiertas por el sistema y si es posible realizarlas en el tiempo y con los conocimientos con los que se cuenta.

La siguiente tabla se muestra cómo se pueden solucionar los problemas descritos en el planteamiento del problema.

Tabla 2 Relación Problema/Solución

PROBLEMA	SOLUCION
Localización de información a fin a las necesidades del usuario	Opción de crear perfil para mejor el filtrado de los resultados de la búsqueda.
Pocas herramientas enfocadas a la cantidad de resultados	El usuario contara con opciones que ayuden al buscador a reducir el número de resultados que regrese
Información repetida o fuentes poco confiables	Registro de fuentes confiables en el perfil de usuario
Resultados de búsqueda mal organizados	Filtrado de información por medio del perfil de usuario
Limitación de búsqueda a palabras claves	Opciones de búsqueda por URL, palabras claves y artículos de Internet

#### 4.1.2 Factibilidad Técnica.

Se cuentan con las herramientas necesarias de software y hardware para realizar el sistema. En cuanto al software se utilizará versiones libres con licencias GNU. En el apartado de Estudio de software se detallara un poco más sobre la utilización de este entorno de desarrollo, al igual que el hardware mostrando las características de la red de computadoras bajo la cual funcionara aplicación.

Tabla 3: Comparación entre servidores

Opción 1	Opción 2
Modelo: Dell PowerEdge R200	Modelo: Toshiba Satellite 550
Procesador: Intel Quad Core Pentium E2200, 2.2Ghz	Procesador: Intel Core 2 Duo
Disco duro: 1 x 160 GB SATA	Disco duro: 500GB SATA
Memoria RAM: 1 GB	Memoria Ram: 4 gB
Taza de Transferencia mensual: 2000 GB	Taza de Transferencia: 4 MB (Dependiente de Conexión a Internet)

### **4.1.3 Estudio de Factibilidad Económica**

El costo para desarrollar del sistema se reduce al costo del equipo de cómputo y el software, para el uso de la licencia del sistema operativo se utilizara una versión de Windows Seven de Microsoft, además, se manejaran librerías de uso libre. En cuanto al equipo de cómputo este se nos ha proporcionado por parte del Instituto Politécnico Nacional, cabe mencionar que se cuenta con el mobiliario, red y hardware necesario.

Por otro lado se necesita de un servidor basado en Apache Tomcat con compilación en Java y de tipo dedicado, las opciones de este tipo de servidores a nivel nacional no son muchas y pocas cumplen con los requerimientos antes mencionados, por ejemplo:

- Nombre del Plan: Servidor Dedicado Linux Mini  
Espacio: 1 x 160 GB SATA  
Transferencia: 2000 GB  
Precio / Año: \$28577.22 MXN
- Nombre del Plan: Servidor Dedicado Linux Económico  
Espacio: 2 x 250 GB SATA  
Transferencia: 5000 GB  
Precio /Año: \$45729.26 MXN

Los precios que se manejan, ambos enlistados son los mas bajos, son infactibles para nosotros el costear el precio de la renta anual de este tipo de servicios, sin mencionar que, debido al periodo que se usaría (6 meses), seria poco practico contratarlo de manera anual, debido a esto se usara una de las computadoras de uno de los miembros del equipo como servidor, debido a que por el momento esta cuenta con los requerimientos necesarios para soportar al numero de usuarios que harán uso del sistema.

### **4.1.4 Estudio de Factibilidad Legal**

Para la realización del proyecto no se usa ningún software que requiera de una licencia de paga, nos sujetaremos a los lineamientos sobre el uso de licencias Express Edition de Microsoft y de GNU GPL (General Public License) o en su defecto de open source.

## **4.3 Análisis de Riesgos**

Valores de impacto:

1. Catastrófico
2. Critico
3. Marginal
4. Despreciable

Tabla 4: Riesgos

Riesgos	Descripción	Tipo	Probabilidad	Impacto
<b>Subestimación del tamaño</b>	El tamaño del proyecto se ha subestimado	Proyecto	Media	2
<b>Mayor número de usuarios de los previstos</b>	Una cantidad muy grandes de usuarios, para los cuales no podría ser suficiente la capacidad de respuesta del servidor.	Proyecto, Técnico y negocio	Media	2
<b>Demasiada información en la base de datos.</b>	Una cantidad muy grandes de usuarios, para los cuales no podría ser suficiente nuestra capacidad de almacenamiento de datos.	Proyecto, Técnico y negocio	Media	2
<b>Reutilización de software</b>	Mala aplicación de técnicas de programación, lo cual no permitiría poder reutilizar código y genera más volumen.	Técnico	Baja	4
<b>Cambio de requerimientos</b>	Habrà más requerimientos de los esperados, o cambios en los requerimientos actuales.	Proyecto y producto	Media	1
<b>La fecha límite está muy ajustada</b>	Subestimar el tiempo de desarrollo provocando retrasos o entregas deficientes.	Proyecto y negocio	Media	1
<b>Bajo rendimiento de las herramientas CASE</b>	Las herramientas CASE que son usadas para el proyecto no tienen el rendimiento esperado	Producto	Baja	4
<b>La tecnología no cumplirá las expectativas</b>	El uso de agentes no es el adecuado para la realización de los filtros adecuados	Técnico	Baja	2
<b>Personal inexperto para el uso de la tecnología</b>	Los desarrolladores no logran implementar de manera eficiente la tecnología necesaria en el sistema debido a su inexperiencia.	Técnico	Media	1
<b>Interactividad con motores de búsqueda</b>	No es posible establecer interacción con el motor de búsqueda seleccionado.	Técnico	Media	2
<b>Sofisticación del usuario</b>	El usuario no logra entender por completo el modo de operar el sistema	Negocio	Baja	3
<b>Documentación del producto</b>	Errores en la documentación del sistema o falta de algún elemento de la misma.	Técnico	Baja	2
<b>Retrasos en la especificación</b>	Las especificaciones de las interfaces esenciales no estarán a tiempo.	Proyecto y producto	Baja	3
<b>Pérdida de algún integrante</b>	Algún integrante del equipo, por cualquier motivo, tiene que abandonar el proyecto	Proyecto	Baja	3
<b>Costo elevado de desarrollo</b>	El desarrollo del sistema tiene costos muy elevados, incosteables para los	Negocio	Baja	1

	integrantes			
<b>Desorganización para el proceso del software</b>	Desentendimiento de la metodología elegida.	Técnico	Alta	2
<b>Escaza información acerca de la tecnología empleada</b>	La documentación y ayuda acerca de la tecnología no es la suficiente para la implementación	Técnica	Baja	3

Tabla 5: Contingencia

Riesgos	Forma de disminuir el riesgo
<b>Subestimación del tamaño</b>	Realizar una investigación a fondo sobre los requisitos e implementación del sistema
<b>Mayor número de usuarios de los previstos</b>	Limitar los usuarios a una cierta región o país
<b>Demasiada información en la base de datos</b>	Emplear técnicas de partición de información
<b>No reutilización de software</b>	Aplicar al máximo la programación orientadas a objetos y algunas otras técnicas.
<b>Cambio de requerimientos</b>	Hacer un estudio profundo sobre los requerimientos, dando el tiempo suficiente para tener todo precisado.
<b>La fecha límite está muy ajustada</b>	Aplicar algún método de gestión de proyectos, como el de la ruta crítica
<b>Bajo rendimiento de las herramientas CASE</b>	Comparar entre las herramientas CASE disponibles, para elegir la más adecuada
<b>La tecnología no satisfará las expectativas</b>	Investigar acerca de otros trabajos que hayan aplicado las tecnologías elegidas, para conocer sus limitantes
<b>Personal inexperto para el uso de la tecnología</b>	Pedir asesorías con expertos en la materia, leer artículos publicados de trabajos similares
<b>Interactividad con motores de búsqueda</b>	Desarrollar un agente especializado, que pueda sustituir al motor de búsqueda
<b>Sofisticación del usuario</b>	Brindar un manual de usuario y crear una interfaz amigable
<b>Documentación del producto</b>	Trabajar en base a las recomendaciones de la metodología
<b>Retrasos en la especificación</b>	Desarrollar la interfaz al tiempo del diseño del sistema
<b>Pérdida de algún integrante</b>	Crear un ambiente cordial de trabajo. Repartir el trabajo de manera que un solo integrante no sea el único encargado de desarrollar alguna tarea.
<b>Costo elevado de desarrollo</b>	Usar software gratuito o realizar un estudio de costo-rendimiento

<b>Desorganización para el proceso del software</b>	Seguir correctamente el calendario de actividades
<b>Escasa información acerca de la tecnología empleada</b>	Buscar en todas las fuentes de información posibles: libros, expertos en la materia, internet, revistas científicas, etc.

#### 4.4 Estudio de software

La complejidad en el diseño de arquitecturas de agentes hace necesario el uso de plataformas de desarrollo que faciliten la programación.

A continuación se muestra una tabla comparativa mostrando algunas características importantes de 4 plataformas que permiten el desarrollo de sistemas basados en arquitectura de agentes.

Tabla 6: Tabla comparativa de Software

	Armazones	Genéricas	FIPA (95)	JAVA	API de agentes	Monitor remoto de agentes	Distribución libre	FIPA-ACL
<b>Zeus</b>	✓			✓	✓		✓	✓
<b>ABLE</b>	✓							
<b>JADEX</b>		✓	✓	✓	✓	✓	✓	✓
<b>Grasshopper</b>		✓						

El entorno elegido para el desarrollo de SURFACE después del estudio, será JADEX, debido a es un entorno robusto y eficiente que soporta todos los servicios básicos de infraestructura especificados, siguiendo los estándares FIPA (comunicaciones, movilidad, gestión de agentes y localización de agentes), a los que añade algunas utilidades gráficas para facilitar la administración de las

JADEX es un paquete desarrollado en Java que cuenta con Java Agent DEvelopment framework, el cual, soporta la implementación de AMS (FIPA compliant):

- Librería de paquetes para programar agentes.
- Herramientas para la depuración y la ejecución de agentes.

Plataforma distribuida para la ejecución de agentes: middle-ware para AMS.

- Servicios para agentes.- directorios, codificación y transporte de mensajes.
- Herramientas.- monitor remoto de agentes, sniffer agent, dummy agent.

## **jQuery**

Es una biblioteca JavaScript que simplifica el recorrido documento HTML, manejo de eventos, animación y las interacciones Ajax para el desarrollo web rápido. Esta diseñado para cambiar la forma en que escribes Java Script.

Proporciona abstracciones de bajo nivel de interacción y animación, efecto avanzados y de alto nivel, widgets temable, construido en la parte superior de la biblioteca jQuery Java Script, que puede utilizar para construir aplicaciones web altamente interactivas.

- Nos ahorra muchas líneas de código.
- Nos hace transparente el soporte de nuestra aplicación para los navegadores principales.
- Nos provee de un mecanismo para la captura de eventos.
- Provee un conjunto de funciones para animar el contenido de la página en forma muy sencilla.
- Integra funcionalidades para trabajar con AJAX

## **Google AJAX**

Google AJAX Search es una API entregada por Google para implementar un buscador utilizando la tecnología AJAX. Este interfaz puede buscar sobre la web, imágenes, blogs y videos. También se le puede restringir a que busque sobre un determinado dominio.

La API de las bibliotecas es una red de distribución de contenidos y la arquitectura de carga para los populares, bibliotecas de Java Script de código abierto.

## **4.5 Arquitectura de Software**

La arquitectura de software se basa en el análisis de las estructuras necesarias para el diseño del mismo. Al principio de la era informática, el grupo de personas capaces de desarrollar código era muy limitado, esto debido a la dificultad que suponía para la mayoría de las personas en ese tiempo. Sin embargo con el paso del tiempo, se fueron desarrollando diferentes metodologías que simplificaban los problemas de la programación de código. A estas metodologías se les llamo “Arquitectura de Software” debido a la similitud que tenía con el desarrollo de planos para realizar una edificación.

## **4.6 Modelo de Arquitectura**

Los modelos de arquitectura se definen dependiendo del tipo de sistema que se esté desarrollando, puesto que cada modelo tiene diferentes características que ayudan a la creación del software.

A continuación mencionaremos algunas de las arquitecturas usadas para el desarrollo de sistemas software:

- **Modelo Estructural.-** En este modelo la arquitectura se muestra como un conjunto organizado de los componentes que conforman el sistema.
- **Modelo Frameworks.-** Este modelo se basa en la búsqueda de patrones arquitectónicos que pueden ser reproducidos en aplicaciones similares.
- **Modelo Dinámico.-** Son modelos dedicados a resaltar los aspectos dinámicos de una arquitectura, indicando como los eventos externos pueden afectar la estructura del sistema.
- **Modelo de 3 capas.-** Este modelo es usado principalmente para diseño de servicios web, puesto que divide en tres módulos principales:
  - Capa de Presentación.- Esta capa contiene la interfaz que interactuará con el usuario.
  - Capa de Lógica o Funcionalidad.- También conocida como “Capa de Negocio”, puesto que contiene la lógica del negocio, en base a que proporciona el puente entre la primera y tercera capa.
  - Capa de Datos.- En esta parte se guardan los datos que contiene la Base de Datos que contiene todos los datos necesarios para el funcionamiento del sistema.
- **Modelo Funcional.-** Son usados para la representación la jerarquía funcional en el sistema.

Tomando en cuenta los requerimientos de nuestro sistema, el modelo de arquitectura que se usara para el desarrollo del sistema será el Modelo de 3 capas, puesto que este modelo se enfoca a servicios web y ofrece las necesidades estructurales que requiere un buscador.

## 4.7 Estilo de Arquitectura

Los estilos de una arquitectura se definen por el tipo de comunicación de datos el cual se pretende desarrollar, por esta causa puede que se confunda con el modelo, sin embargo el modelo es previo al desarrollo del sistema y los estilos se asignan dependiendo del sistema que se pretende desarrollar.

Los estilos de arquitectura describen, como se puede ver en el listado de la parte de abajo, que orientación tengan, a continuación mencionaremos algunos de estos:

- **Centrado en Datos.-** Como parte central de este estilo aparece un almacén de datos, el cual tiene como principal característica, que es ingresado de manera frecuente por los usuarios de sistema.
- **Centrado en Flujo de Datos.-** Basada en patrones de “pipe & filter”, tuberías y filtros, consta de un conjunto conexiones entre filtros bidireccionales, que pueden tanto recibir como enviar información.
- **Llamada y respuesta.-** Este estilo tiene como principal característica el permitir a los desarrolladores crear estructuras que puedan modificar y escalar de manera fácil. Podemos encontrar diferentes tipos dentro de este estilo:
  - Programa principal.- Se descomponen las funciones de modo jerárquico donde un programa principal puede llamar a un programa subordinado, y este puede invocar a su vez a otros.
  - Llamadas de procedimiento remoto.- Los diferentes componentes de un sistema se encuentran distribuidos entre diferentes computadoras.
- **Orientada a objetos.-** El sistema se divide en bloques los cuales encapsulan datos y operaciones que utilizan para manipular dichos datos
- **Capas.-** Definido por un conjunto de niveles o capas, cada nivel interno que se atraviesa se aproxima más al nivel del conjunto de instrucciones. Cada capa solo puede comunicarse con las vecinas, facilitando así la portabilidad del diseño.

Cada diseño ofrece ventajas y cualidades al sistema donde se utilice. Dependiendo de las necesidades que el sistema necesita, se usara el estilo de “llamada y respuesta” puesto que nuestro sistema utiliza subprogramas para lograr tanto la búsqueda como la aplicación de los filtros.

## 4.8 Metodología de Desarrollo Web

Aplicación web es un sistema el cual consulta, procesa y estructura los datos mediante navegadores. Las características de este tipo de sistemas son que la interacción con el usuario es muy alta además que el diseño de la interfaz debe de ser claro, simple de tal manera que oriente a cualquier tipo de usuario en nuestro caso usuarios registrados o no.

La propuesta de ingeniería web basada en UML (Koch, 2000) es una metodología detallada para el proceso de autoría de aplicaciones con definición exhaustiva del proceso de diseño que debe de ser utilizado.

Los principales aspectos son:

- Uso de una notación estándar para todos los modelos (UML Lenguaje de modelo unificado)
- Definición de métodos (Los pasos para la construcción de los modelos)
  - FASES DEL DESARROLLO WEB
  - En proceso de la aplicación está dividido en cuatro pasos o actividades
- Análisis de requisitos. Fija los requisitos fundamentales de la aplicación reflejados en un modelo de casos de uso
- Diseño conceptual. Materializado en un modelo de dominio considerando los requisitos reflejados en los casos de uso
- Diseño Navegaciones
  - Modelo del espacio de navegaciones
  - Modelo de la estructura de navegación.
- Diseño de presentación. Representa las vistas de la interfaz del usuario mediante modelos estándares de interacción UML

## Capítulo 5 Diseño del sistema

### 5.1 Requerimientos Funcionales y No Funcionales

#### Requerimientos Funcionales

- El Sistema deberá realizar búsquedas a partir de palabras clave en la *World Wide Web* mediante un navegador Web.
- El sistema deberá permitir el Registro de Perfiles de Usuario.
- El Sistema deberá permitir la Modificación de Perfiles de usuario previamente registrados.
- El sistema deberá permitir la Eliminación de Perfiles de usuario previamente registrados.
- El sistema deberá permitir el registro de “Fuentes Confiables” dentro del Perfil de Usuario previamente creado.
- El sistema deberá permitir la modificación y eliminación de “Fuentes confiables” dentro del Perfil de Usuario previamente creado.
- El sistema podrá ser utilizado tanto por usuarios que Registren un Perfil de Usuario, como por usuario que no se registren.

#### Requerimientos no Funcionales

- Para crear un perfil de usuario será necesario proporcionar:
  - Nombre Completo del usuario.
  - Correo Electrónico (no previamente registrado).
- En el perfil de usuario se podrán registrar datos conocidos como “Fuentes Confiables”. Dichas fuentes confiables constarán de Autores de Artículos de confianza, Páginas Web de preferencia (URL) y temas en específico. Esta información será de utilidad para el filtro de los resultados realizado por la estructura de Agentes Web.
- Las búsquedas a través de internet se realizarán de la siguiente manera dependiendo del tipo de usuario:
  - Usuario no Registrado: El usuario proporcionará las palabras claves con las cuales desea realizar la búsqueda. Se realizará un filtro de la información tomando en cuenta el historial de navegación del usuario (excluyendo páginas innecesarias a través del chequeo de los metadatos de las mismas) y las palabras claves proporcionadas. Dicha búsqueda y

filtrado de la información será realizado por la estructura de Agentes Web, tomando en cuenta aspectos como el historial del usuario, las palabras clave.

- Usuario Registrado: El usuario proporcionará las palabras claves con las cuales desea realizar la búsqueda. Se realizará un filtro de la información tomando en cuenta el historial de navegación del usuario (excluyendo páginas innecesarias a través del chequeo de los metadatos de las mismas), el Perfil de Usuario (Búsquedas personalizadas), y las palabras claves proporcionadas. Dicha búsqueda y filtrado de la información será realizado por la estructura de Agentes Web, tomando en cuenta aspectos como el historial del usuario, las palabras clave, las preferencias registradas, y el perfil personalizado de búsqueda que se esté utilizando, dándole preferencia posicional en la tabla de resultados a los que estén relacionados por el Perfil de Búsqueda Personalizada utilizado en ese momento.
- El Sistema podrá ser utilizable desde el navegador Web de preferencia del usuario (Firefox, Google Chrome, Internet Explorer, etc.), incorporando los datos de navegación del usuario (como cookies e historial) para los resultados de búsqueda.
- Para su acceso, el servidor deberá estar montado sobre un Host y tener un dominio asignado.

## 5.2 Diagrama a bloques

El los diagramas a bloques son representaciones graficas del cómo funciona internamente un sistema. En el diagrama a bloques que se muestra a continuación, se puede observar el funcionamiento del sistema, pasando por su Navegador, después por Internet y finalmente accediendo al sistema el cual se encontrara almacenado en un host.

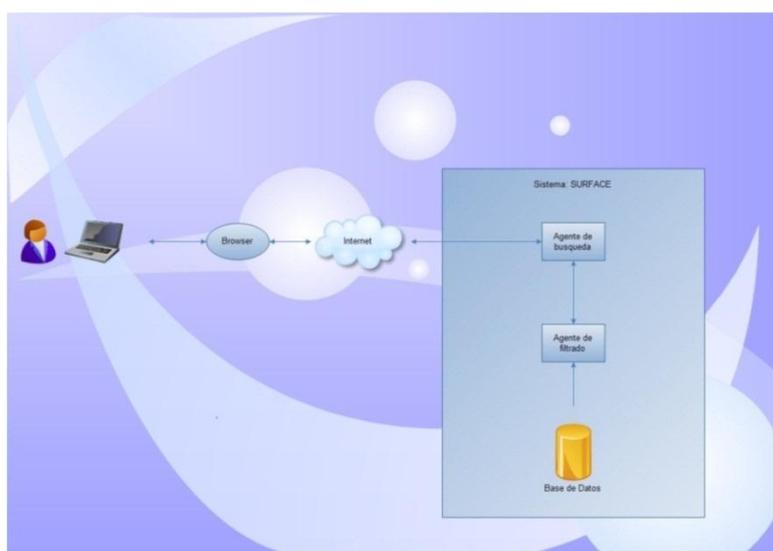


Figura 3: Diagrama a Bloques

### 5.3 Definición de casos de uso

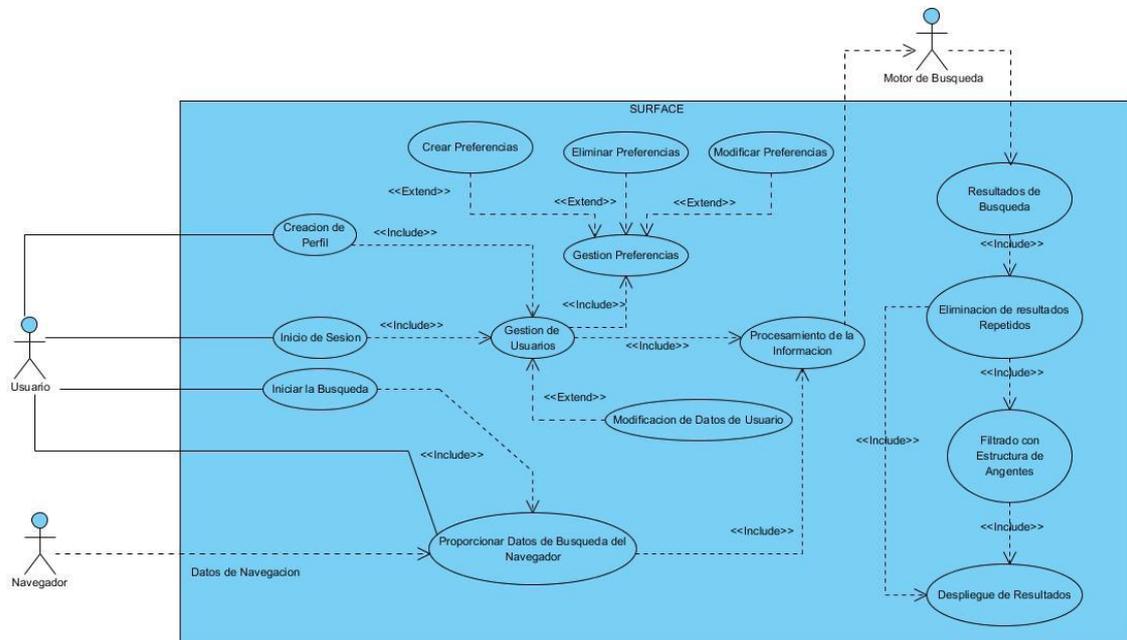


Figura 4: Diagrama de Casos de Uso

ID	CU1
Nombre	Creación de perfil
Descripción	El sistema mostrara un formulario, el cual deberá ser contestado por el usuario, para así poder realizar búsquedas que utilicen criterios especializados para esta.
Objetivo	Obtener datos que ayudaran al sistema a proporcionar al usuario resultados que sean a fines con sus preferencias personales.
Actores	Usuario.
Entradas	
Salidas	<ul style="list-style-type: none"> <li>Perfil de usuario.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>El usuario debe haber solicitado la creación de un perfil.</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>El usuario accesa al sistema desde de Navegador.</li> <li>Fin.</li> </ol>

ID	CU 2
Nombre	Inicio de Sesión
Descripción	El sistema le concede acceso a un usuario que tiene un previo registro en el mismo.
Objetivo	El usuario que cuenta con un perfil en el sistema utilizara este a su máximo potencial, puesto que el perfil proporciona un filtro que ayudara al sistema a entregar resultados más precisos.
Actores	Usuario
Entradas	<ul style="list-style-type: none"> <li>• Id. Usuario</li> <li>• Contraseña.</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Autorización de acceso al sistema.</li> <li>• Carga de perfil de usuario.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario debe tener un perfil previo creado, para que pueda acceder a la búsqueda.</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario introduce su Id. Usuario.</li> <li>2. El usuario introduce su contraseña.</li> <li>3. El sistema le da acceso al usuario.</li> <li>4. Fin.</li> </ol>

ID	CU 3
Nombre	Iniciar Búsqueda
Descripción	Usuario con o sin cuenta introduce sus parámetros de la búsqueda que realizara.
Objetivo	El sistema deberá regresar resultados relacionados a las palabras clave que el usuario utilice en el campo de búsqueda.
Actores	Usuario.
Entradas	<ul style="list-style-type: none"> <li>• Palabras clave.</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Palabras clave.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario debe haber entrado al sistema.</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario introduce sus palabras clave en el campo de búsqueda.</li> <li>2. Las palabras clave son enviadas al procesador de información de búsqueda.</li> <li>3. Fin</li> </ol>

ID	CU 4
Nombre	Gestionar Usuarios
Descripción	La gestión de usuarios permite la creación y uso de los perfiles necesarios para definir las búsquedas personalizadas.
Objetivo	La gestión de los usuarios permite al usuario la creación y modificación de su perfil y definir las búsquedas personalizadas.
Actores	Usuario.
Entradas	<ul style="list-style-type: none"> <li>• Datos de perfil preestablecido.</li> <li>• Datos de nuevo perfil.</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Datos de perfil de usuario.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario debe haber creado un perfil.</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario solicita entrar a su perfil. <ol style="list-style-type: none"> <li>a. El perfil ya existe.</li> <li>b. El perfil aún no ha sido creado.</li> </ol> </li> <li>2. Fin.</li> </ol>

ID	CU 5
Nombre	Modificación de Datos de Usuario
Descripción	El sistema permite a los usuarios tanto la creación y modificación de sus datos personales en el sistema.
Objetivo	El usuario debe poder ser capaz de crear o modificar su perfil. Actores: Usuario.
Actores	Usuario
Entradas	<ul style="list-style-type: none"> <li>• Datos del perfil de usuario</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Perfil actualizado.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario debe tener un perfil.</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario accede al sistema. <ol style="list-style-type: none"> <li>a. El usuario crea su perfil.</li> <li>b. El usuario modifica su perfil.</li> </ol> </li> <li>2. Fin</li> </ol>

ID	CU 6
Nombre	Gestionar Preferencias.
Descripción	El sistema da acceso al usuario a un perfil de búsqueda nuevo y en caso de tener uno preestablecido la modificación o eliminación del mismo.
Objetivo	El sistema debe poder permitir al usuario la modificación, creación y eliminación de sus preferencias para la búsqueda personalizada.
Actores	Usuario.
Entradas	<ul style="list-style-type: none"> <li>• Direcciones URL</li> <li>• Temas Concretos.</li> <li>• Autores.</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Perfil de Búsqueda Actualizado.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario debe tener una cuenta</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario accede al sistema.</li> <li>2. El usuario selecciona la opción de perfil de búsqueda. <ol style="list-style-type: none"> <li>a. El usuario crea un perfil de búsqueda.</li> <li>b. El usuario modifica un perfil de búsqueda. <ol style="list-style-type: none"> <li>i. El usuario modifica la tabla de autores.</li> <li>ii. El usuario modifica la tabla de URL.</li> <li>iii. El usuario modifica la tabla de Temas.</li> </ol> </li> <li>c. El usuario elimina un perfil de búsqueda.</li> </ol> </li> <li>3. Fin</li> </ol>

ID	CU 7
Nombre	Crear Preferencias
Descripción	El usuario crea un perfil de búsqueda
Objetivo	Las búsquedas personalizadas requieren de un perfil de búsqueda de usuario, que sumado a los parámetros de la búsqueda actual, se utilizara para realizar búsquedas por perfil de usuario
Actores	Usuario
Entradas	<ul style="list-style-type: none"> <li>• Nuevo perfil de usuario. <ul style="list-style-type: none"> <li>○ Nuevo Autor.</li> <li>○ Nuevo URL.</li> </ul> </li> </ul>
Salidas	Nuevo Tema. <ul style="list-style-type: none"> <li>• Perfil de búsqueda actualizado</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• Usuario cuenta con registro en el sistema.</li> <li>• Inicio de sesión con el usuario.</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario inicia sesión.</li> <li>2. El usuario selecciona la opción de gestión de perfiles de búsqueda. <ol style="list-style-type: none"> <li>a. El usuario crea un perfil.</li> <li>b. El usuario introduce al menos un dato por campo.</li> </ol> </li> <li>3. Fin.</li> </ol>

ID	CU 8
Nombre	Eliminar Preferencias
Descripción	El usuario puede no requerir más un perfil, así que se le dará la opción de eliminarlo
Objetivo	El eliminar un perfil de búsqueda que ya no es requerido por el usuario.
Actores	Usuario
Entradas	<ul style="list-style-type: none"> <li>• Autorización de eliminación de perfil</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Perfil eliminado</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario debe haber iniciado una sesión.</li> <li>• El usuario debe contar con al menos un perfil de búsqueda</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario inicia sesión.</li> <li>2. El usuario selecciona la opción de gestión de perfiles de búsqueda.</li> <li>3. El usuario elimina un perfil.</li> <li>4. Fin.</li> </ol>

ID	CU 9
Nombre	Modificar Preferencias
Descripción	El usuario modifica un perfil existente
Objetivo	El usuario podrá ser capaz de modificar un perfil preexistente en su cuenta.
Actores	Usuario
Entradas	<ul style="list-style-type: none"> <li>• Nuevos parámetros de búsqueda o modificación de los mismos. <ul style="list-style-type: none"> <li>○ Modificación de tabal de autores.</li> <li>○ Modificación de tabla de URL.</li> <li>○ Modificación de tabla de temas</li> </ul> </li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Perfil de búsqueda actualizado.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario debe haber iniciado sesión.</li> <li>• El usuario debe contar con al menos un perfil de búsqueda.</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario inicia sesión.</li> <li>2. El usuario selecciona la opción de gestión de perfil de búsqueda.</li> <li>3. El usuario modifica un perfil existente. <ol style="list-style-type: none"> <li>a. El usuario modifica la tabla de autores.</li> <li>b. El usuario modifica la tabla de URL.</li> <li>c. El usuario modifica la tabla de temas.</li> </ol> </li> <li>4. Fin.</li> </ol>

ID	CU 10
Nombre	Procesamiento de Información de Búsqueda
Descripción	El sistema reúne la información de búsqueda y la prepara para ser enviada al motor de búsqueda
Objetivo	Enviar un paquete de información al motor para que este pueda realizar una búsqueda
Actores	<ul style="list-style-type: none"> <li>• Motor de Búsquedas</li> </ul>
Entradas	<ul style="list-style-type: none"> <li>• Parámetros de búsqueda actual.</li> <li>• Perfil de búsqueda de usuario.</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Compilado de datos de búsqueda</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• El usuario haya iniciado una búsqueda</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario accesa al sistema. <ol style="list-style-type: none"> <li>a. El usuario comienza una búsqueda.</li> <li>b. El usuario inicia sesión. <ol style="list-style-type: none"> <li>i. El usuario comienza una búsqueda.</li> </ol> </li> </ol> </li> <li>2. El usuario introduce las palabras clave de la búsqueda nueva.</li> <li>3. El sistema compila la información de búsqueda y del navegador para ser enviado al motor.</li> <li>4. Se envía paquete de información al motor.</li> <li>5. Fin.</li> </ol>

ID	CU 11
Nombre	Obtención de Resultados
Descripción	El motor regresa los resultados de la búsqueda realizada en internet para que se muestren al usuario o se procese con un agente de filtrado
Objetivo	Tener los resultados necesarios para el tratamiento de la información de la búsqueda, para que el sistema pueda procesar la información ya sea para un usuario con registro o sin registro
Actores	<ul style="list-style-type: none"> <li>• Motor de Búsqueda</li> </ul>
Entradas	<ul style="list-style-type: none"> <li>• Resultados de la Búsqueda en Internet</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Resultados de Búsqueda en Internet</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• Búsqueda realizada.</li> <li>• Resultados del Motor</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario accesa al sistema. <ol style="list-style-type: none"> <li>a. El usuario comienza una búsqueda. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> <li>b. El usuario inicia sesión. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> </ol> </li> <li>2. El sistema solicita el historial de navegación.</li> <li>3. El sistema reúne la información de navegación, los parámetros de la nueva búsqueda, en caso de existir, el sistema recibe el perfil de búsqueda del usuario.</li> <li>4. El sistema envía paquete de datos de información al motor de búsqueda.</li> <li>5. El motor realiza la búsqueda y regresa resultados.</li> <li>6. El sistema obtiene los resultados.</li> <li>7. Fin.</li> </ol>

ID	CU 12
Nombre	Eliminación de Resultados Repetidos
Descripción	El sistema toma los resultados del motor, y los analiza para eliminar posibles resultados que se encuentren más de una vez.
Objetivo	Eliminar resultados cuyas fuentes sean las mismas y así hacer mas fácil el filtrado y/o despliegue de los datos, así como evitar que el usuario repita resultados de búsqueda
Actores	
Entradas	<ul style="list-style-type: none"> <li>• Resultados de búsqueda</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Resultados de búsqueda sin repeticiones.</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• Búsqueda realizada.</li> <li>• Resultados del Motor</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario accesa al sistema. <ol style="list-style-type: none"> <li>a. El usuario comienza una búsqueda. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> <li>b. El usuario inicia sesión. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> </ol> </li> <li>2. El sistema solicita el historial de navegación.</li> <li>3. El sistema reúne la información de navegación, los parámetros de la nueva búsqueda, en caso de existir, el sistema recibe el perfil de búsqueda del usuario.</li> <li>4. El sistema envía paquete de datos de información al motor de búsqueda.</li> <li>5. El motor realiza la búsqueda y regresa resultados.</li> <li>6. El sistema obtiene los resultados.</li> <li>7. El sistema eliminar resultados repetidos.</li> <li>8. Fin.</li> </ol>

ID	CU 13
Nombre	Filtración con Estructura de Agentes
Descripción	El sistema recibe los resultados de la búsqueda y los ordena conforme a la puntuación que se les da dependiendo de las similitudes que tenga con los parámetros de búsqueda y preferencias del usuario
Objetivo	Usar la estructura de agentes para dar un filtrado basado en las preferencias del usuario
Actores	
Entradas	<ul style="list-style-type: none"> <li>• Resultados de búsqueda sin repeticiones</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Resultados de búsqueda ordenados</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• Búsqueda realizada.</li> <li>• Resultados del Motor</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario accede al sistema. <ol style="list-style-type: none"> <li>a. El usuario comienza una búsqueda. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> <li>b. El usuario inicia sesión. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> </ol> </li> <li>2. El sistema solicita el historial de navegación.</li> <li>3. El sistema reúne la información de navegación, los parámetros de la nueva búsqueda, en caso de existir, el sistema recibe el perfil de búsqueda del usuario.</li> <li>4. El sistema envía paquete de datos de información al motor de búsqueda.</li> <li>5. El motor realiza la búsqueda y regresa resultados.</li> <li>6. El sistema obtiene los resultados.</li> <li>7. El sistema elimina resultados repetidos.</li> <li>8. El sistema ordena los resultados.</li> <li>9. Fin.</li> </ol>

ID	CU 14
Nombre	Despliegue de Resultados
Descripción	El sistema muestra en pantalla los resultados de la búsqueda realizada, dependiendo del tipo de usuario, el sistema mostrara los resultados solo sin repeticiones o con los resultados organizados dependiendo de las preferencias del usuario y elimina resultados repetidos.
Objetivo	Mostrar al usuario en pantalla los resultados de su búsqueda, para que así, pueda visitar los sitios y recabar su información.
Actores	
Entradas	<ul style="list-style-type: none"> <li>• Resultados de Búsqueda</li> </ul>
Salidas	<ul style="list-style-type: none"> <li>• Resultados de Búsqueda Filtrados</li> </ul>
Precondiciones	<ul style="list-style-type: none"> <li>• Búsqueda realizada.</li> <li>• Resultados del Motor.</li> <li>• Resultados de Filtrado</li> </ul>
Trayectoria principal	<ol style="list-style-type: none"> <li>1. El usuario accesa al sistema. <ol style="list-style-type: none"> <li>a. El usuario comienza una búsqueda. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> <li>b. El usuario inicia sesión. <ol style="list-style-type: none"> <li>i. El usuario introduce los parámetros de la búsqueda actual.</li> </ol> </li> </ol> </li> <li>2. El sistema solicita el historial de navegación.</li> <li>3. El sistema reúne la información de navegación, los parámetros de la nueva búsqueda, en caso de existir, el sistema recibe el perfil de búsqueda del usuario.</li> <li>4. El sistema envía paquete de datos de información al motor de búsqueda.</li> <li>5. El motor realiza la búsqueda y regresa resultados.</li> <li>6. El sistema obtiene los resultados.</li> <li>7. El sistema eliminar resultados repetidos</li> </ol>

## 5.4 Diagrama de Búsqueda y Filtrado

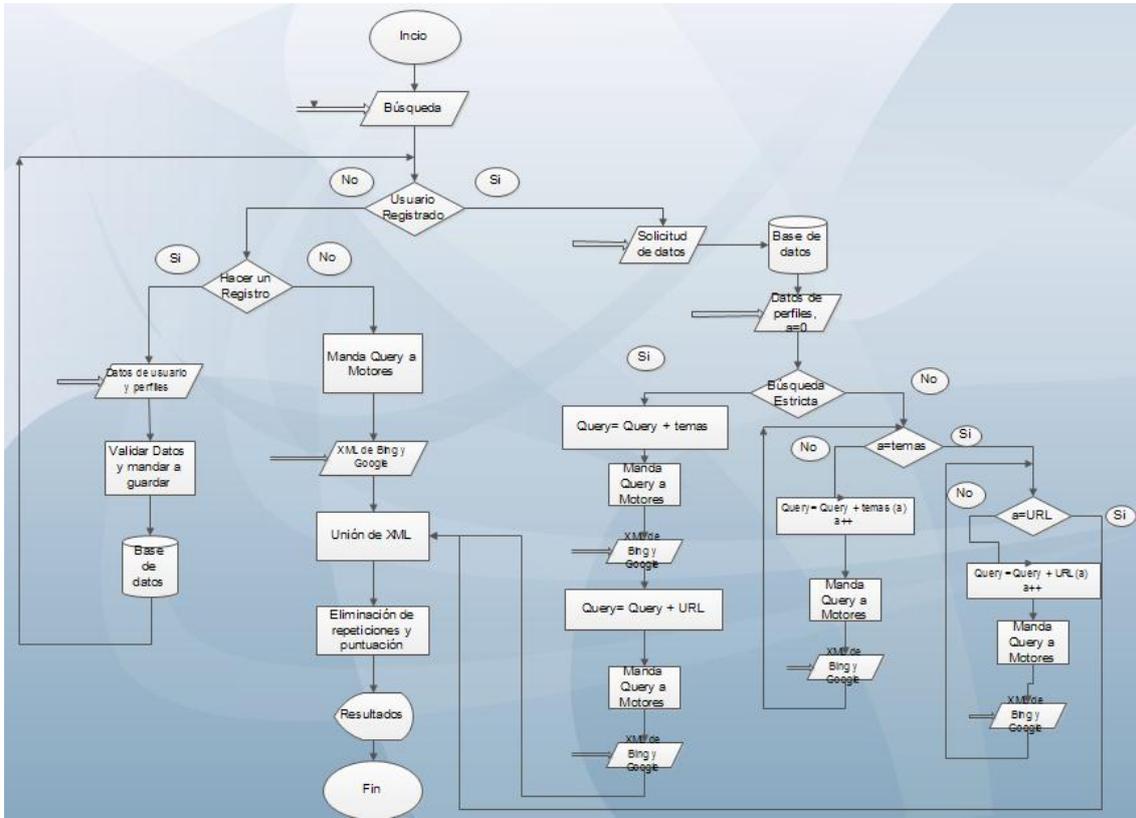


Figura 5: Diagrama de Búsqueda y Filtrado

## 5.5 Diagramas de Secuencia

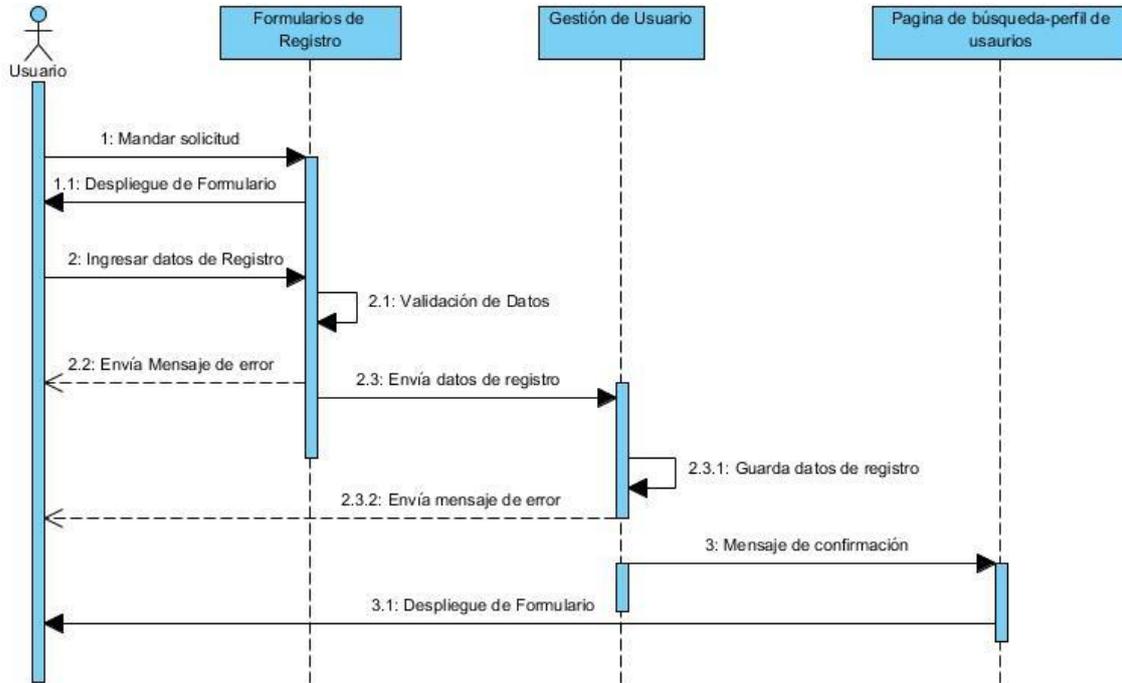


Figura 6: Diagrama de Secuencia Crear Perfil

En el diagrama de creación de perfil muestra como un usuario puede crear una cuenta en SURFACE. Esto llenando los datos que se le piden en el formulario de registro. Estos son nombre, correo, contraseña, genero, país. El sistema validara los datos y si es correcto lo enviara a la base de datos en la cual se guardarán para cuando se inicie la sesión.

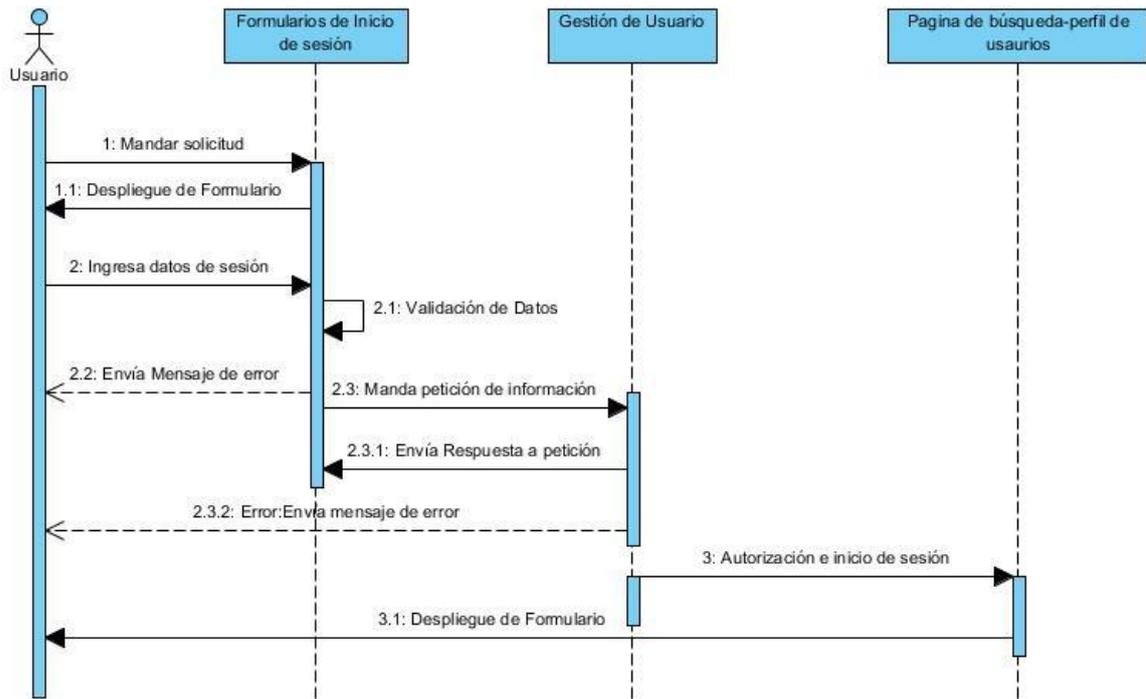


Figura 7: Diagrama de Secuencia: Iniciar Sesión.

Para que los usuarios registrados puedan disfrutar de los beneficios que le ofrece SURFACE tienen que acceder al sistema esto mediante el correo y la contraseña que ingreso en su registro, esto por medio de un formulario. El sistema validara los datos y sacara los registros de la base de datos, comparara los datos ingresados y los de la base de datos y si es correcto ingresara al los formularios de perfil y búsqueda.

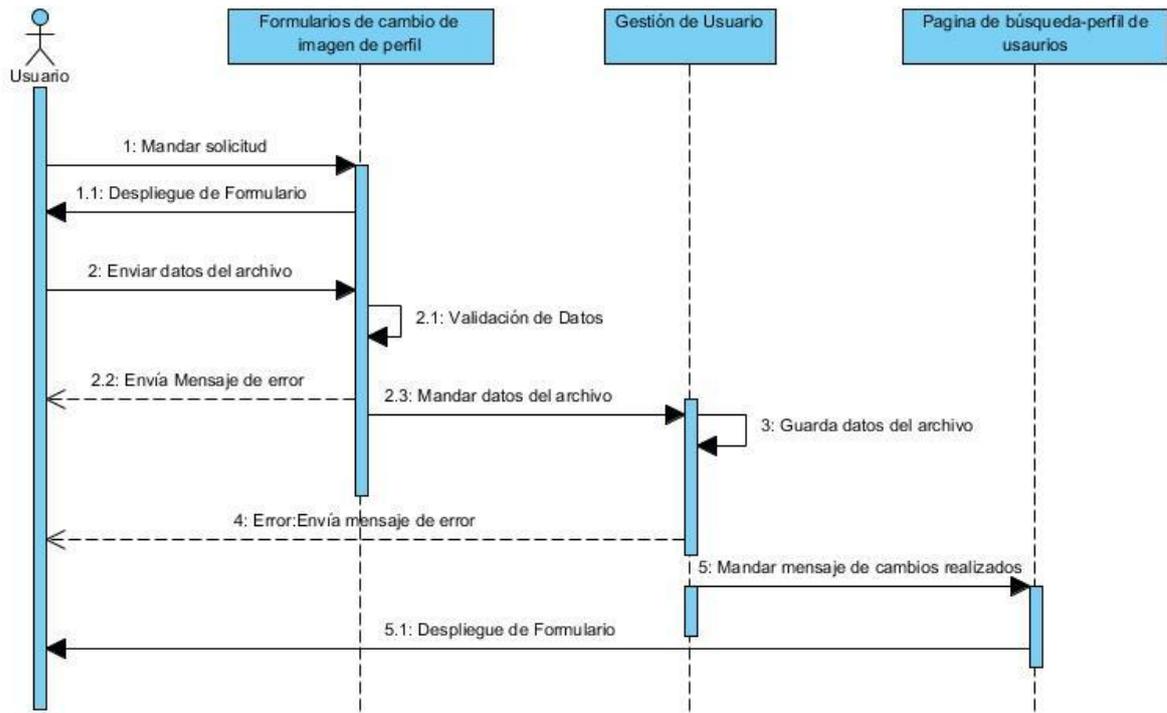


Figura 8: Diagrama de Secuencia: Modificar imagen de perfil.

Para el cambio de la imagen de perfil, previamente se inicia sesión.

Da clic en la imagen para que pueda seleccionar la imagen deseada, se validara el tipo de archivo y el tamaño de la imagen el cual si es correcto se guarda en la base de datos y se manda un mensaje de que los cambios han sido realizado. Y se re direcciona al formulario de perfil y búsqueda del usuario.

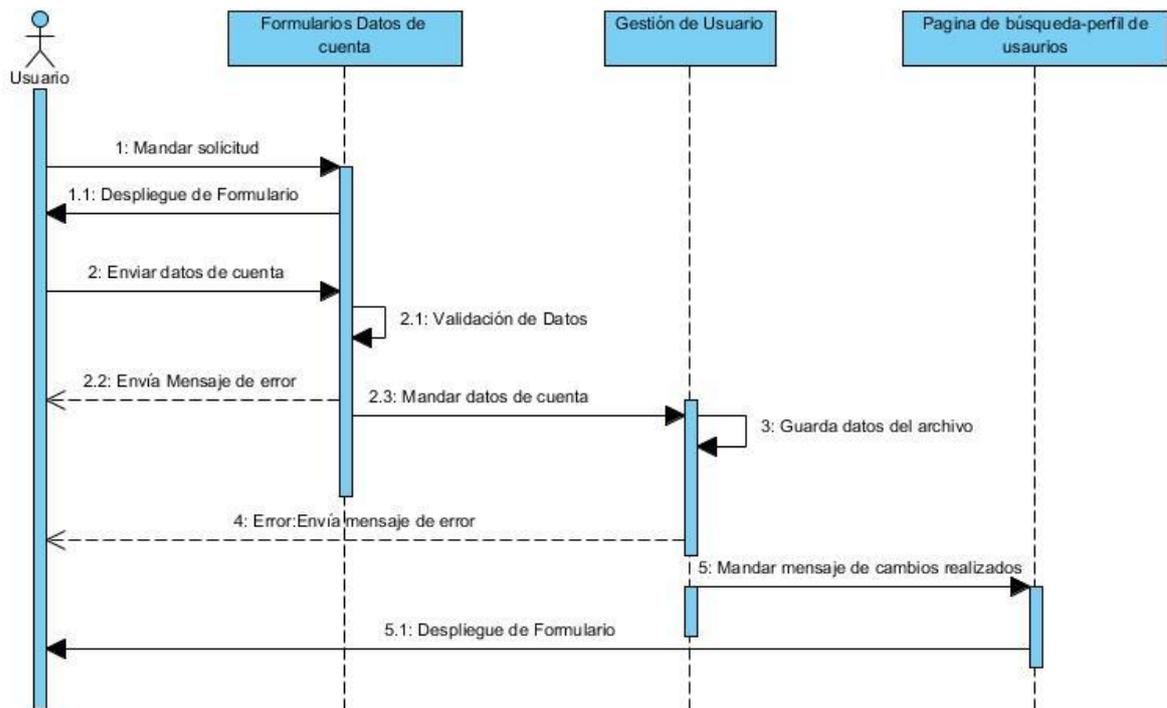


Figura 9: Diagrama de Secuencia: Modificar datos de cuenta.

Previamente el usuario tiene que iniciar sesión, en la modificación de datos de cuenta consta de cambiar la contraseña con la que entra a SURFACE.

El sistema validara la contraseña actual, la nueva contraseña y la confirmación de esta última. Si todo es correcto mandara la nueva contraseña a la base de datos la cual la guardara y mandara un mensaje que se ha realizado el cambio satisfactoriamente.

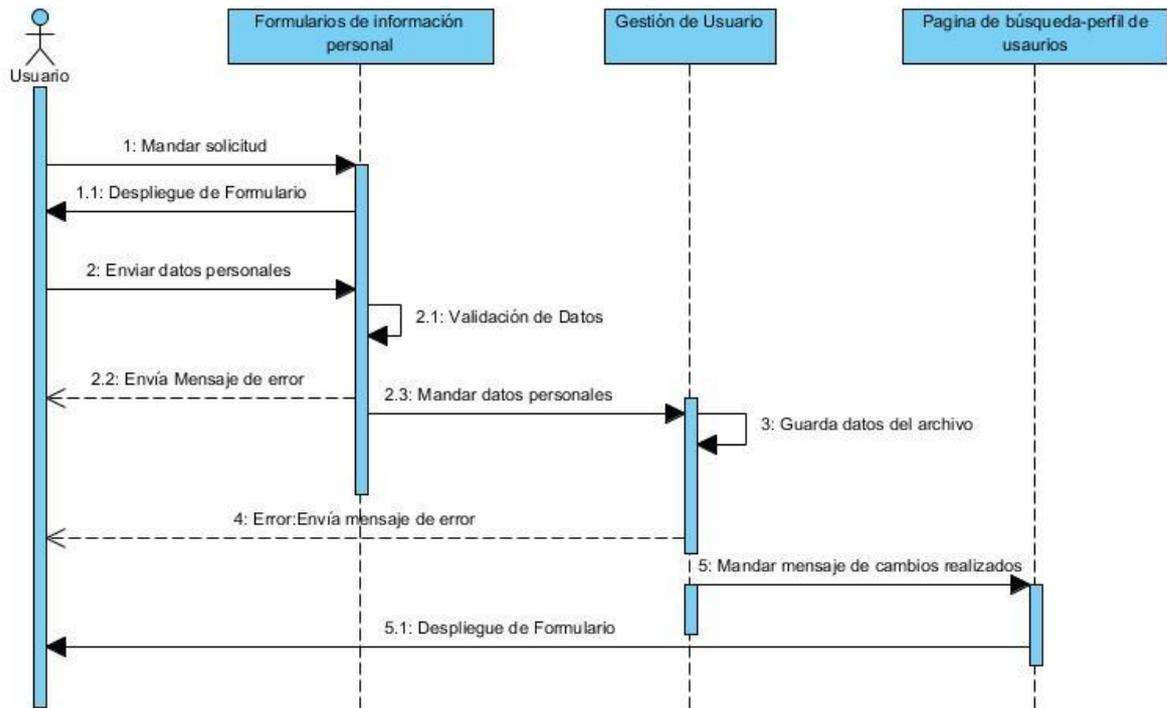


Figura 10: Diagrama de Secuencia: Modificar datos personales

Previamente el usuario iniciara sesión, en el cual el formulario de perfil y búsqueda del usuario podrá acceder para cambiar su información personal esta es nombre, fecha de nacimiento, país y sexo. El sistema validara los datos ingresados y si estos son correctos lo mandara a la base de datos para guardarlos. Si todo es correcto se mandara un mensaje de cambios realizados.

En dado caso que haya un error se mandara el mensaje de error.

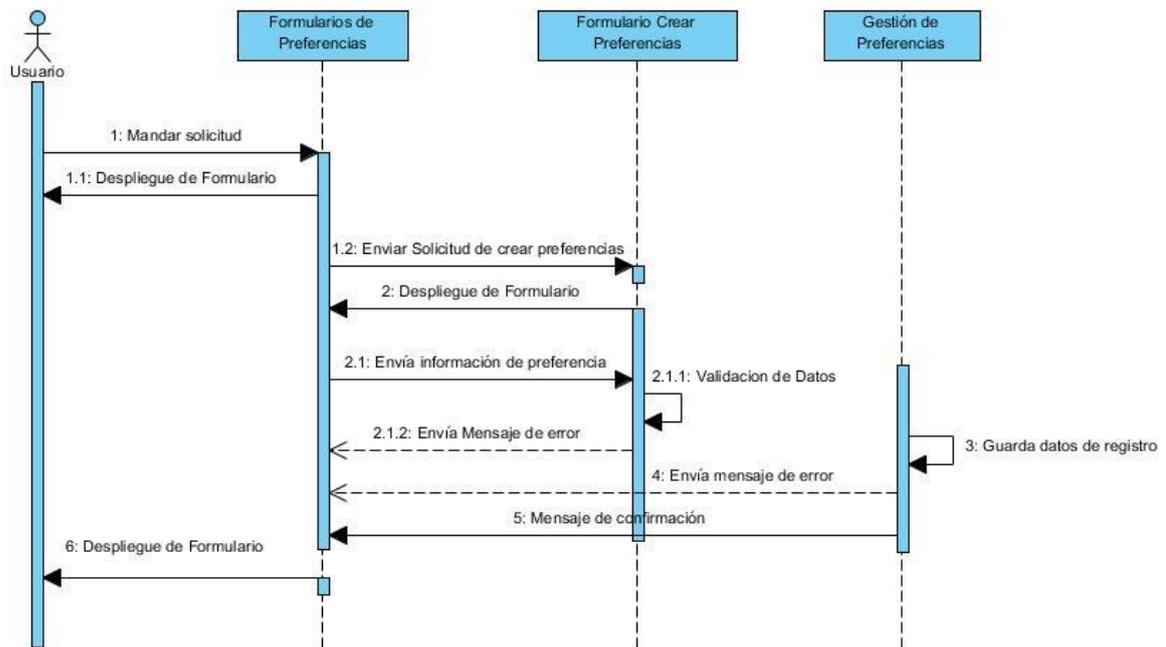


Figura 11: Diagrama de Secuencia: Crear perfil personalizado de búsqueda

Previamente se iniciara sesión y el usuario seleccionara el crear un perfil de búsqueda. El usuario mandara a la información del perfil, el sistema validara los datos, en caso de ser correcto se enviara a la base de datos para guardarlos y enviar un mensaje de perfil guardado.

En caso de tener un error, se enviara un mensaje de error.

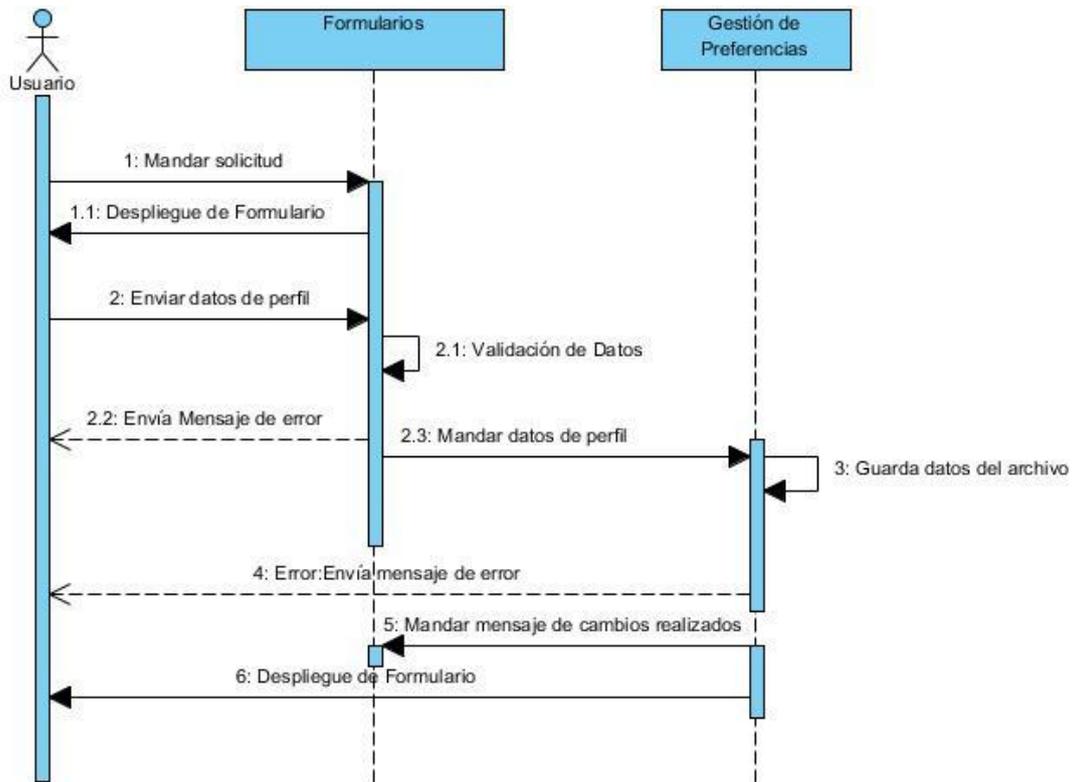


Figura 12: Diagrama de Secuencia: Modificar perfil personalizado de búsqueda

Se inicia sesión para el acceso a la modificación del perfil personalizado de búsqueda. Al acceder, el usuario enviara los datos nuevos del perfil, el sistema validara los datos y al ser correctos se modificaran en la base de datos. Este enviara un mensaje de datos de perfil guardados. En caso contrario se enviara un mensaje de error.

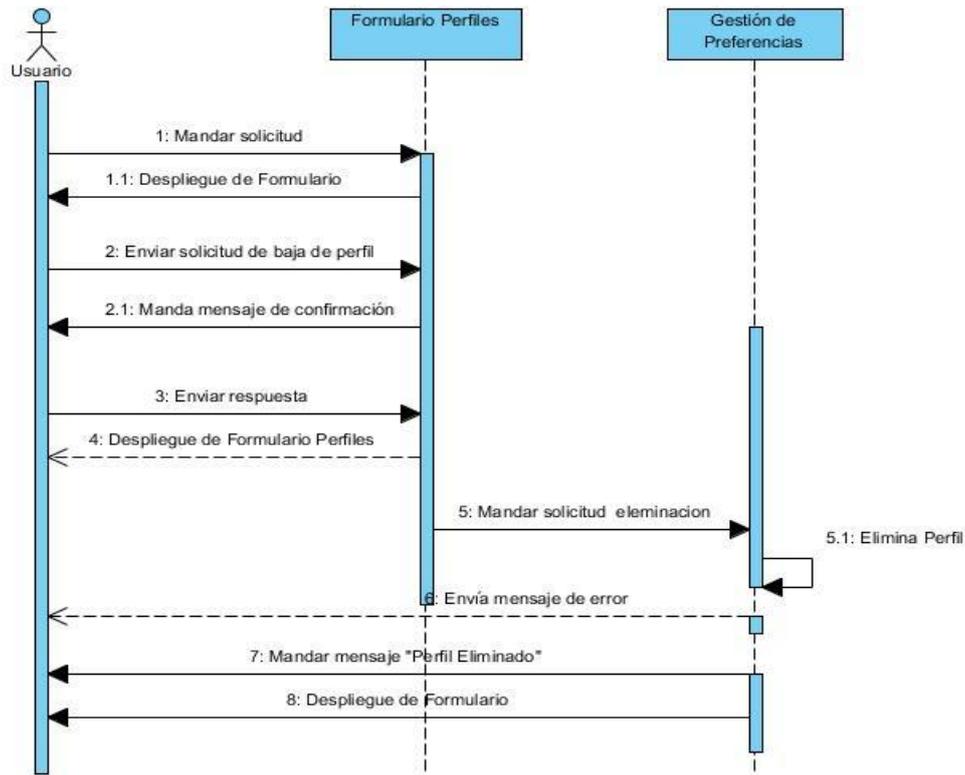


Figura 13: Diagrama de Secuencia: Eliminar perfil personalizado de búsqueda

Para eliminar un perfil personalizado es necesario iniciar sesión, al ingresar al formulario podrá eliminar los perfiles. El sistema mandara un mensaje de confirmación si el usuario desea eliminar, al ser correcto, el perfil será eliminado de la base de datos y mandara un mensaje de perfil eliminado.

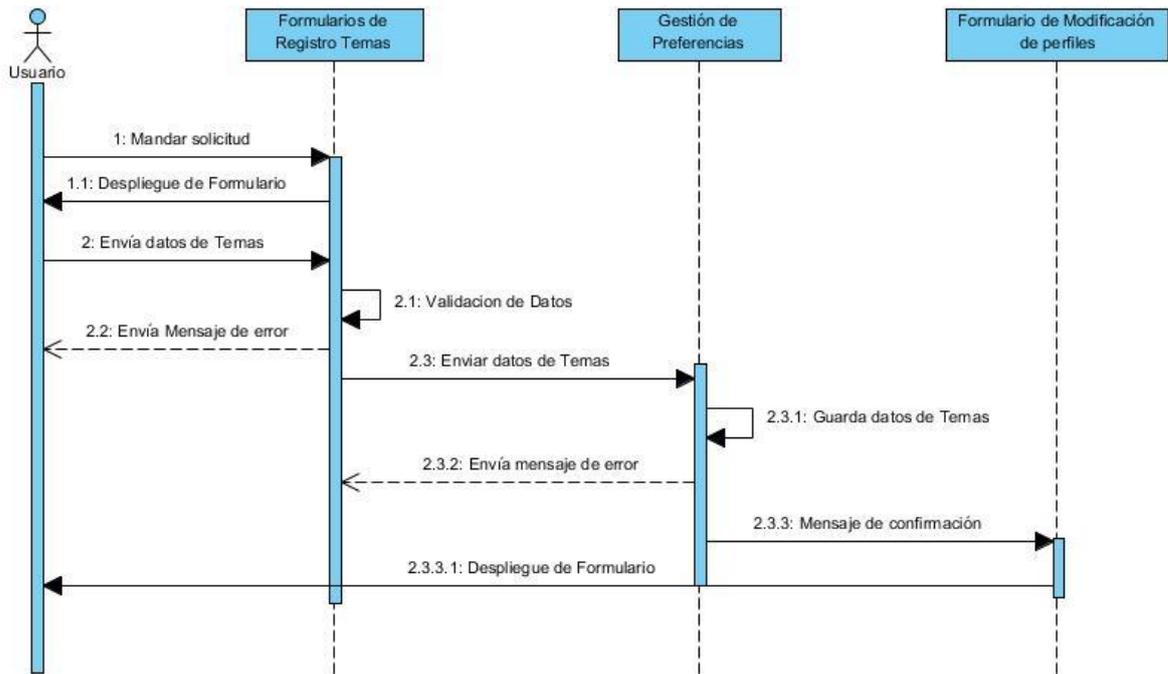


Figura 14: Diagrama de Secuencia: Registrar temas

Previamente se inicia sesión, para el registro de temas, se accede a la modificación de perfiles y al registro de temas, el usuario envía datos del tema que desea registrar. El sistema validara los datos ingresados, en caso de estar correcto estos datos los guarda la base de datos y envía un mensaje de temas guardados. Y redirige a la modificación de perfiles.

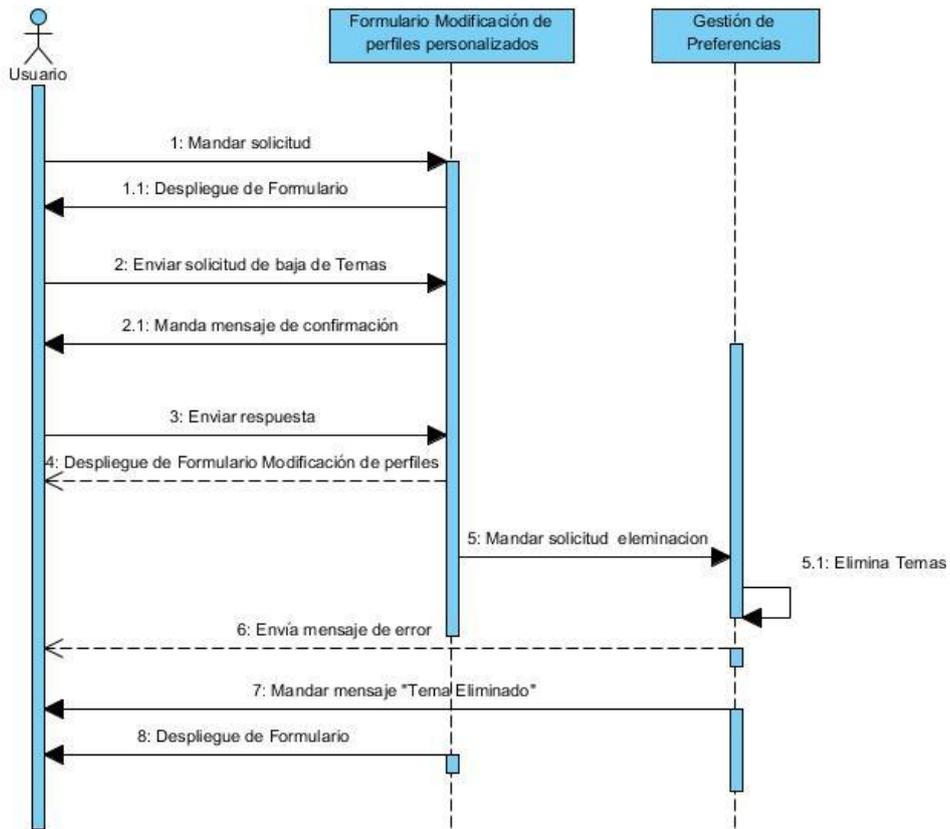


Figura 15: Diagrama de Secuencia: Eliminación temas

Previamente el usuario iniciara sesión, se accede a la modificación de perfiles. El usuario mandara una solicitud de eliminación de temas. El sistema mandara un mensaje de confirmación al ser afirmativo, la base de datos eliminara el tema y mandara un mensaje de tema eliminado. Se regresara al formulario de modificaciones de perfiles.

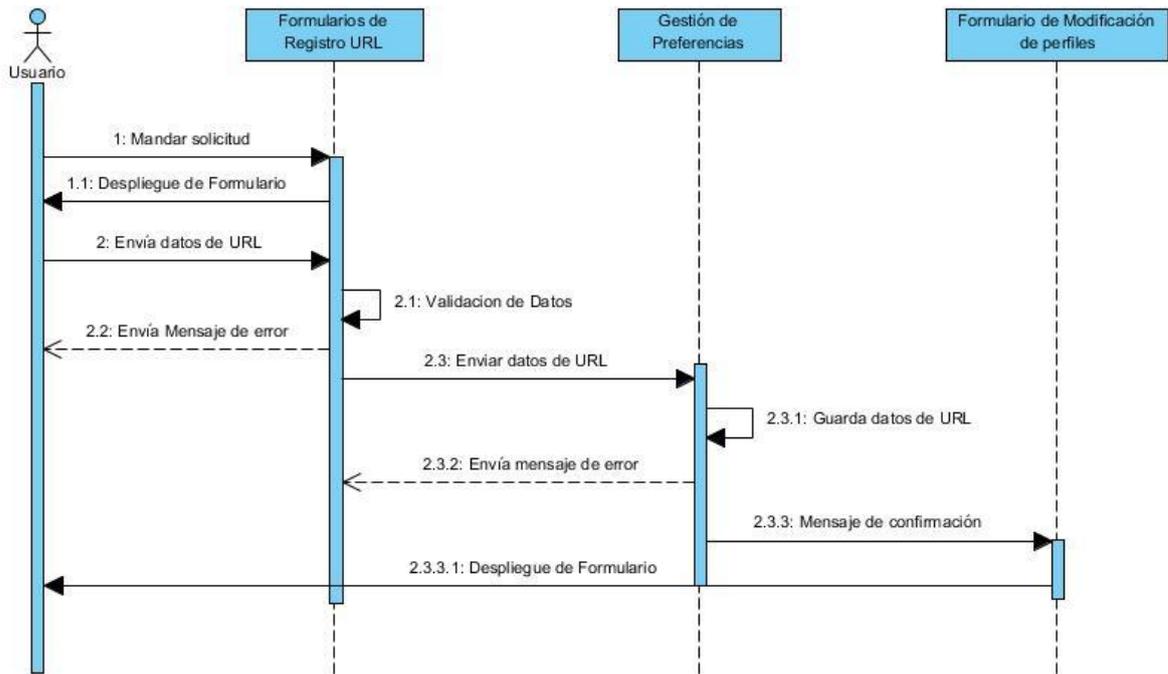


Figura 16: Diagrama de Secuencia: Registrar URL

Previamente se inicia sesión, se accede a la modificación de perfiles y al registro de URL, el usuario envía datos de la URL que desea registrar.

El sistema validara los datos ingresados, en caso de estar correcto estos datos los guarda la base de datos y envía un mensaje de temas guardados. Y redirige a la modificación de perfiles.

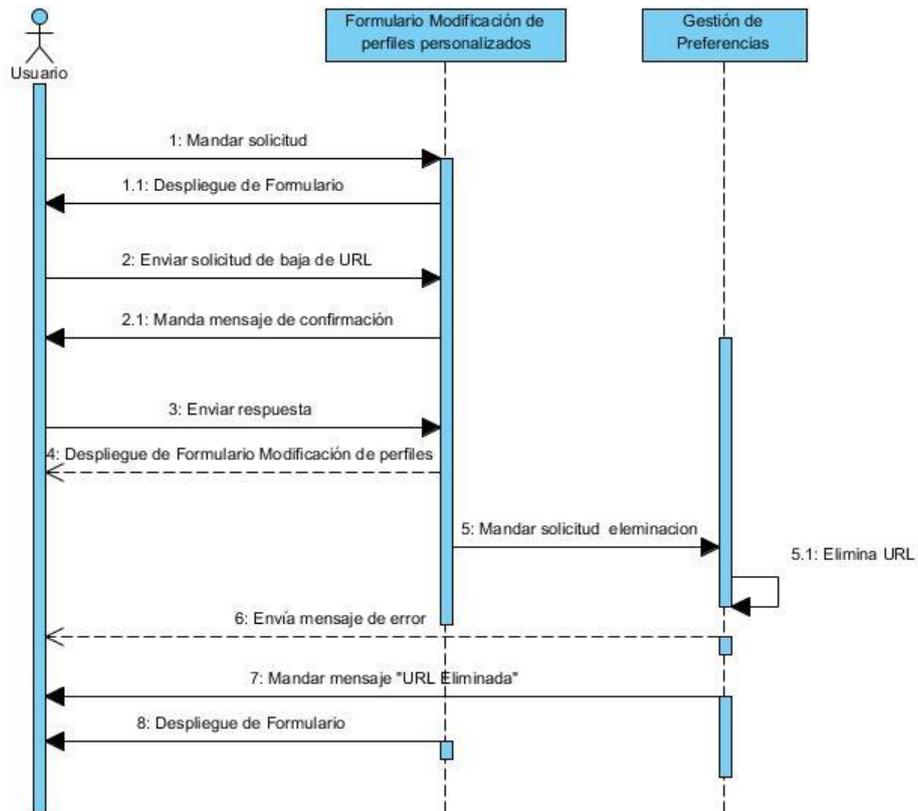


Figura 17: Diagrama de Secuencia: Eliminación URL

Previamente el usuario iniciara sesión, se accede a la modificación de perfiles. El usuario mandara una solicitud de eliminación de URL. El sistema mandara un mensaje de confirmación al ser afirmativo, la base de datos eliminara la URL y mandara un mensaje de URL eliminado. Se regresara al formulario de modificaciones de perfiles.

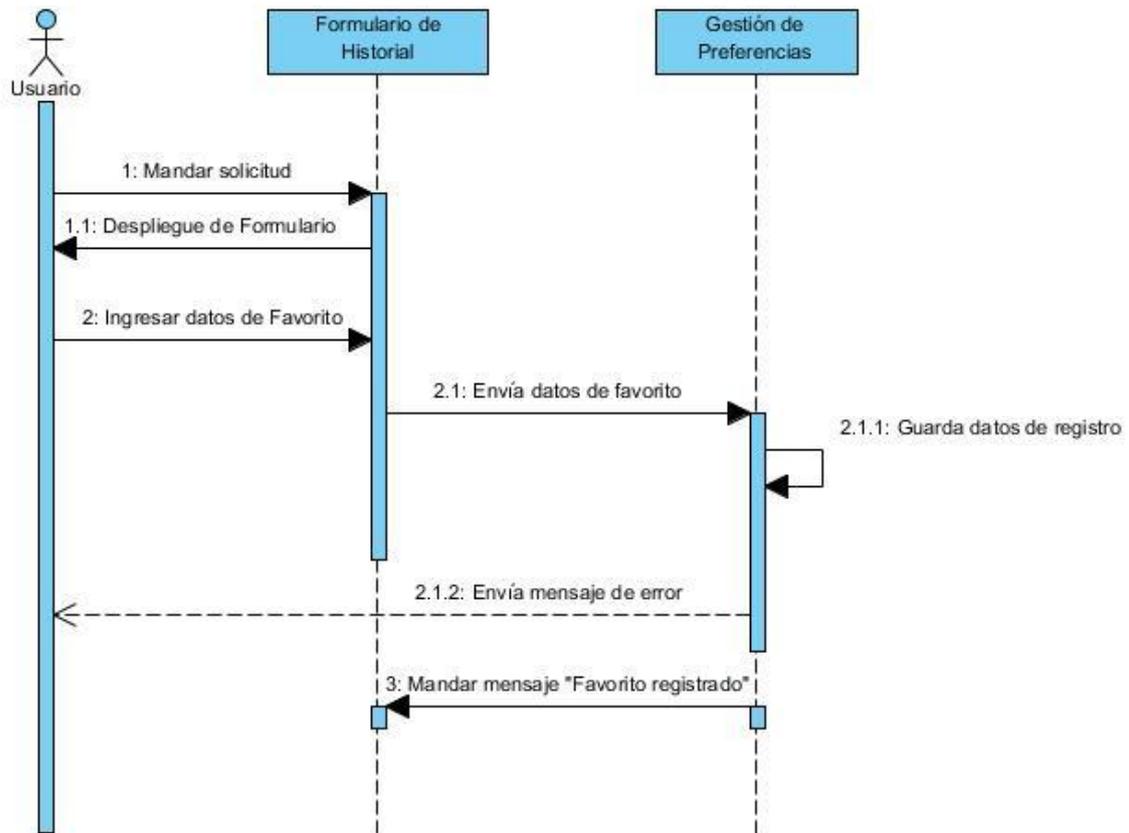


Figura 18: Diagrama de Secuencia: Registro de Favoritos

Previamente se inicia sesión, se accede al formulario de historial. En donde se dará de alta los Favoritos que son algunos resultados que ha sido resultado de búsquedas anteriormente realizadas y que las fuentes le han parecido interesante y/o buenas fuentes de información.

El sistema guardara en la base de datos los favoritos he enviara un mensaje de favorito registrado.

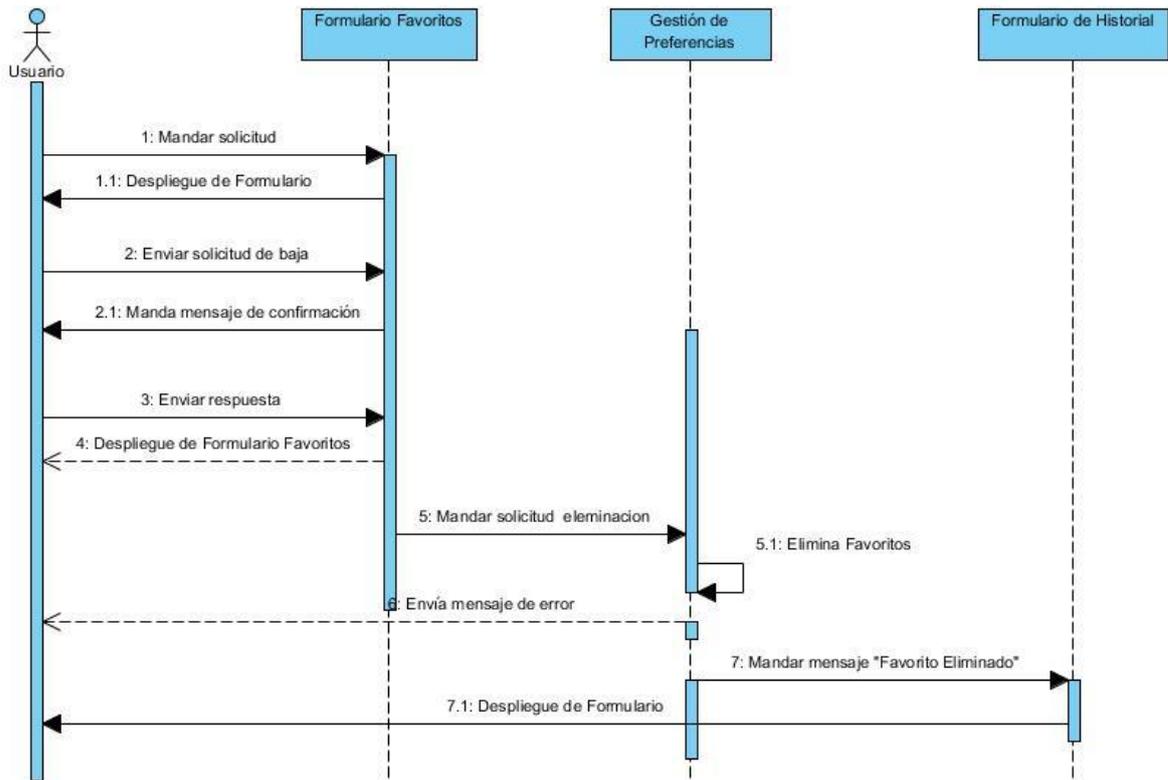


Figura 19: Diagrama de Secuencia: Eliminación de Favoritos

Previamente se inicia sesión, se accede al formulario de favoritos. El usuario mandara una solicitud de baja de favoritos. El sistema mandara un mensaje de confirmación y si es afirmativa la respuesta del usuario se mandara la eliminación en la base de datos los favoritos seleccionados, he enviara un mensaje de favorito eliminado.

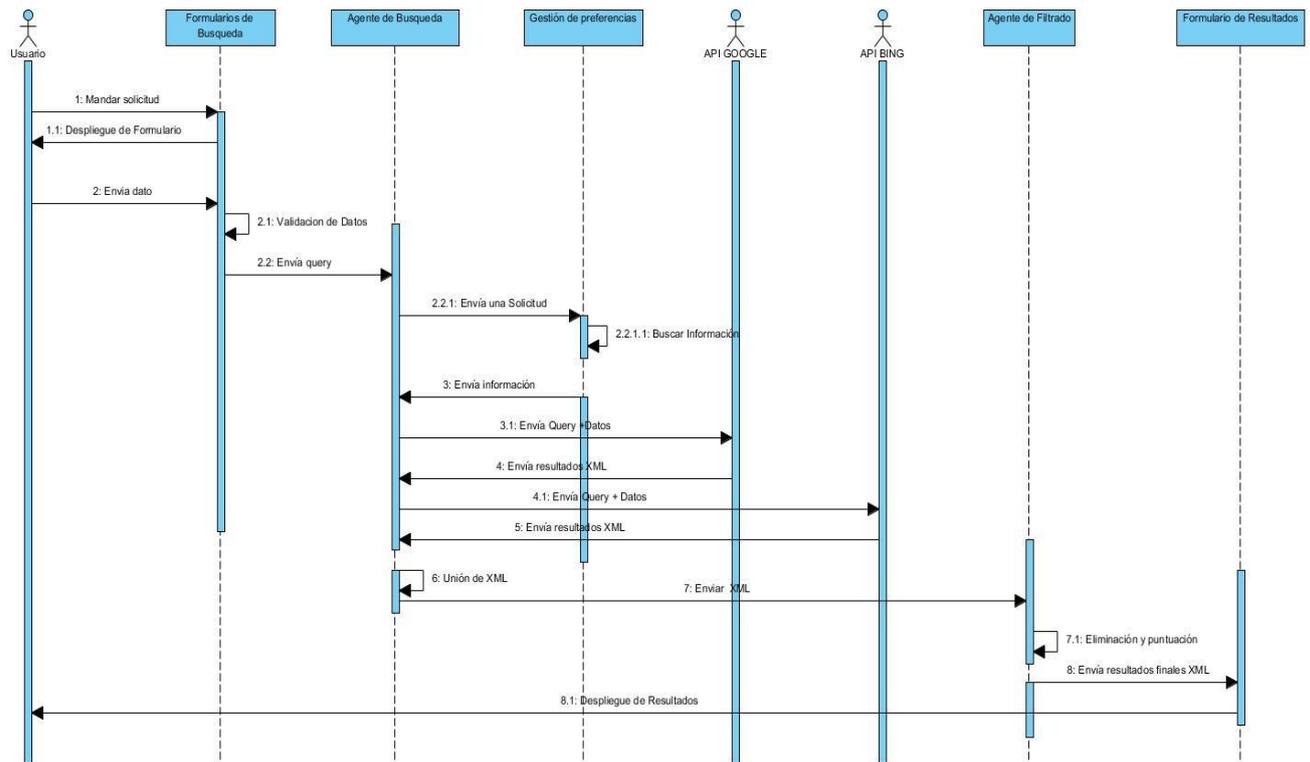


Figura 20: Diagrama de Secuencia: Realizar búsqueda con registro

Para realizar una búsqueda normal es necesario estar registrado por lo que el usuario podrá disfrutar de los beneficios de SURFACE si ingresa a su sesión.

El usuario ingresara la búsqueda a realizar y seleccionara el perfil que desea utilizar, el sistema validara los datos y si es correcto enviara un Query y la información del perfil al agente de búsqueda este se encargara de enviar una solicitud de información a la base de datos el cual si es correcto enviara la información de temas y URL almacenados en el perfil.

El motor enviara el Query mas un autor o temas a los motores de búsqueda BING y GOOGLE a estos a su vez realizaran la búsqueda y le regresara una lista de resultados al agente de búsqueda, así sucesivamente hasta terminar con los temas y URL almacenados.

El agente unirá los resultados y posteriormente los mandara al agente de filtrado este es el encargado de eliminar repeticiones y de a su vez poner una puntuación a los resultados. Los resultados serán puestos en un formulario de resultados y serán desplegados al usuario.

Entiéndase como puntuación, al proceso para decidir el orden de despliegue de los resultados obtenidos a partir de una búsqueda, se maneja un sistema de puntos: Por cada resultado duplicado se eliminara su repetición y se le asignara un "punto". Esto se realizara con todos los resultados; al final, se colocaran en la parte superior aquellos que tengan mayor puntuación (de manera descendente).

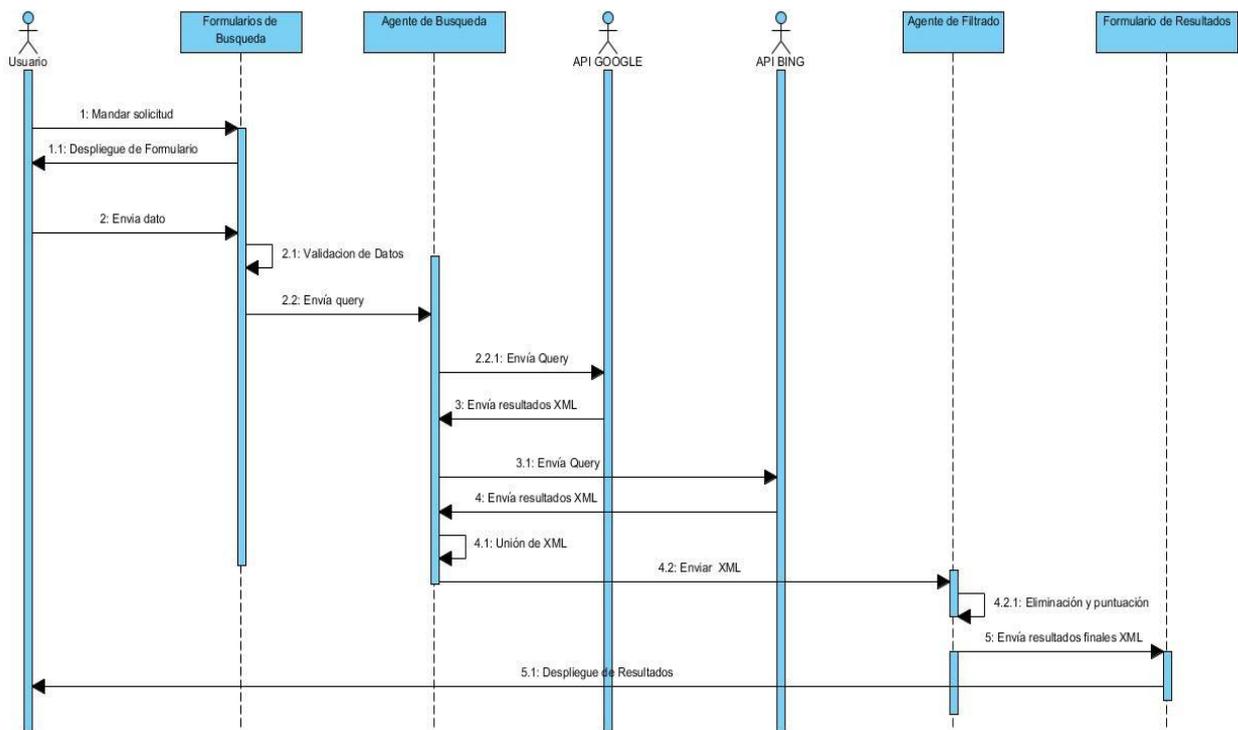


Figura 21: Diagrama de Secuencia: Realizar búsqueda sin registro

Para realizar una búsqueda normal no es necesario estar registrado por lo que el usuario podrá realizar una consulta.

El usuario ingresara la búsqueda a realizar el sistema validara los datos y si es correcto enviara un Query al agente de búsqueda este se encargara de enviar ese Query a los motores de búsqueda BING y GOOGLE a estos a su vez realizaran la búsqueda y le regresara una lista de resultados al agente de búsqueda, este unirá los resultados y posteriormente los mandara al agente de filtrado este es el encargado de eliminar repeticiones y de a su vez poner una puntuación a los resultados. Los resultados serán puestos en un formulario de resultados y serán desplegados al usuario.

Entiéndase como puntuación, al proceso para decidir el orden de despliegue de los resultados obtenidos a partir de una búsqueda, se maneja un sistema de puntos: Por cada resultado duplicado se eliminara su repetición y se le asignara un "punto". Esto se realizara con todos los resultados; al final, se colocaran en la parte superior aquellos que tengan mayor puntuación (de manera descendente).

## 5.6 Diagrama de Actividades

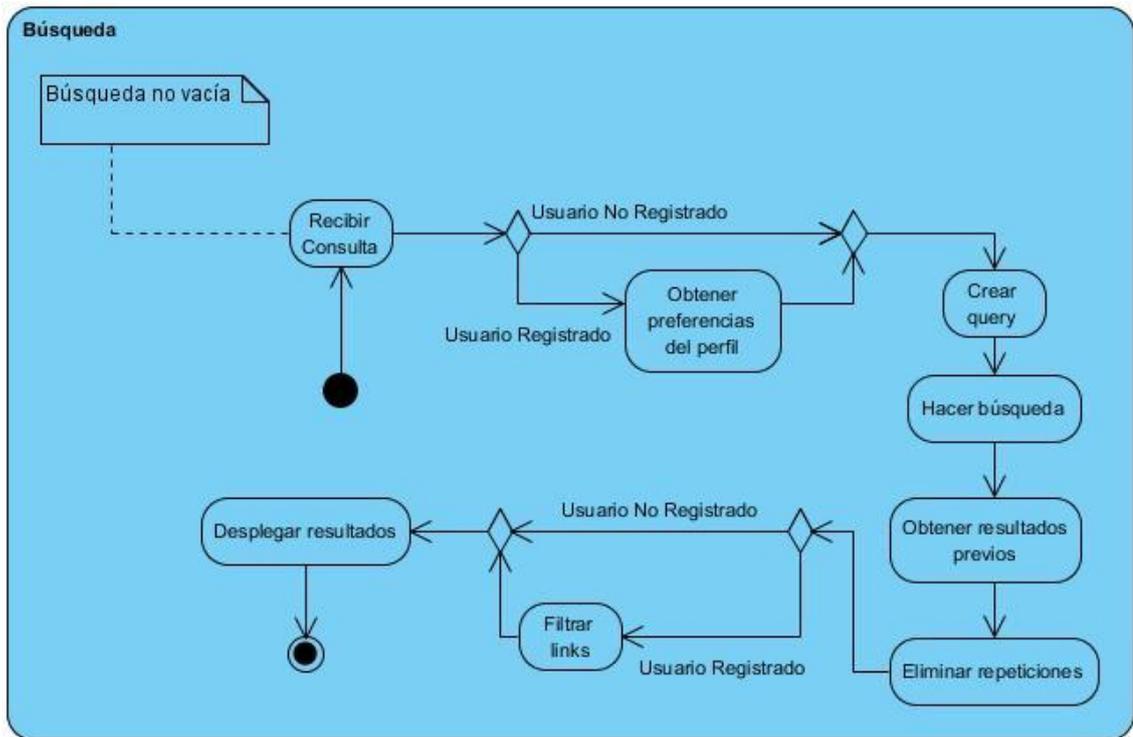


Figura 22: Diagrama de Actividades

Cuando se realiza una búsqueda (no vacía), se recibe la consulta del usuario, enseguida hay dos opciones: la primera es que la búsqueda la esté realizando un usuario registrado, y la segunda es que quien la realiza sea un usuario no registrado. En la primer opción, se accesa a la base de datos para obtener las preferencias del usuario, con éstos datos se genera la consulta final que se envía a los motores de búsqueda, en caso del usuario no registrado, la consulta del usuario se pasa directamente a los motores.

Posteriormente, obtenemos los resultados que arrojan los motores, se combinan y eliminamos repeticiones de links, y enseguida, solo si se trata de un usuario registrado, se filtran los links según sus preferencias.

Finalmente, los resultados son desplegados en el navegador.

## 5.7 Diagrama de Estado

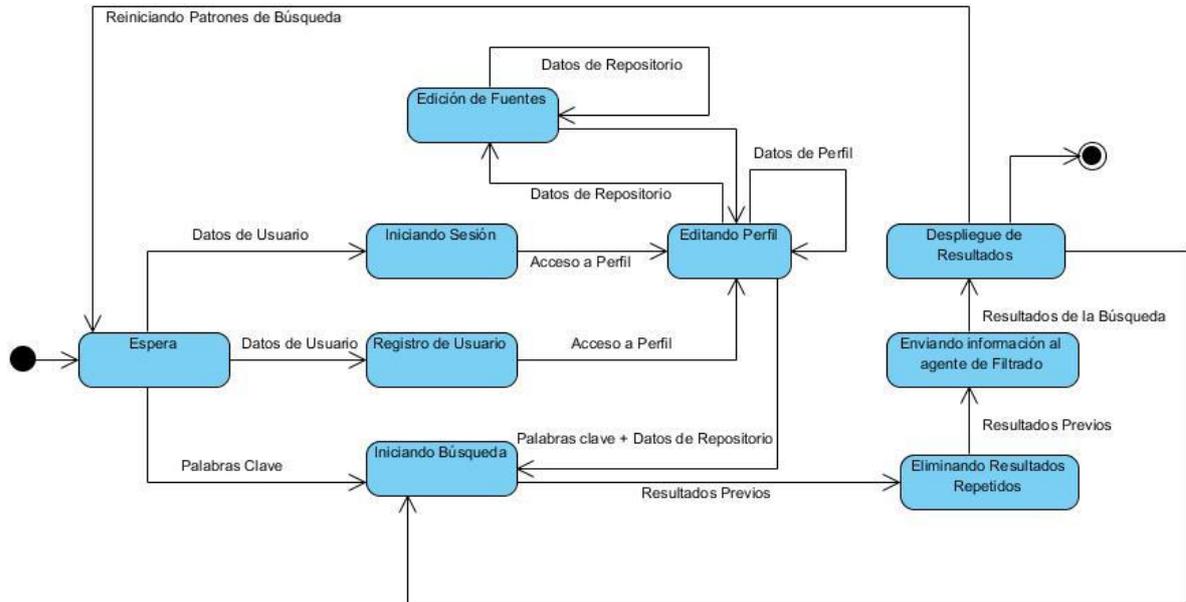


Figura 23: Diagrama de Estados

El sistema comienza en un estado de espera y puede pasar a alguno de los tres siguientes estados: Iniciando sesión, registro de usuario e iniciando búsqueda.

Los estados iniciando sesión y registro de usuario pueden pasar al estado editando sesión, donde es posible editar los, y editando las fuentes que sería ahora el estado de edición de fuentes. Editando perfil puede pasar también al estado iniciando búsqueda, que se explica enseguida.

El estado iniciando búsqueda, donde se pueden acceder palabras clave para empezar a realizar consultas, la consulta se realiza y los resultados devueltos por los motores se manipulan, primeramente, se eliminan resultados repetidos, estos a su vez, se envían al agente para ser filtrados según las preferencias del usuario y finalmente se despliegan los resultados. Después de cualquier proceso el sistema vuelve a un estado de espera.

## 5.8 Diagrama Relacional

El diseño de la base de datos está elaborado cumpliendo hasta un tercer nivel de normalización separando la información lo más posible para evitar la repetición innecesaria de la información y realizar búsquedas ágiles en menor tiempo.

Descripción de la Base de Datos:

Se manejará un catálogo de cuentas, usuarios (perfiles), un catálogo de Países y de Imágenes.

Toda esta información en conjunto conformará al usuario como tal. El conjunto de los ID de dichos catálogos se formará en una tabla de detalle.

Se tomaron en cuenta las siguientes características:

- Una cuenta puede tener solo 1 Detalle de Usuario, pasando su ID
- Un país se puede encontrar en varias tuplas de detalle, pasando su ID
- Una misma imagen puede estar en varias tuplas de detalle, pasando su ID
- La información Personal (Perfil) del usuario, puede estar en solo 1 tupla del detalle, pasando su ID

Se manejará un catálogo de Tipos de Perfil (para las Búsquedas Personalizadas), el conjunto de los ID del catálogo de Tipos de Perfil e Información del Usuario se formará en una tabla de detalle, dando origen a todos los tipos de Perfil que cree un usuario.

Se tomaron en cuenta las siguientes características:

- Un usuario (Perfil de usuario) puede tener varios detalles de Tipos de Perfil, pasando el ID de Usuario
- Un tipo de Perfil puede tener solo 1 Detalle de Tipos de Perfil, pasando el ID del Tipo de Perfil

Por último, cada tipo de perfil contiene un número determinado de Autores, URL y Temas relacionados al Tipo de Perfil que ha creado el usuario, la información se organizará en tablas de Detalle las cuales tendrán las siguientes características:

- Un Tipo de Perfil puede estar contenido en varias tuplas de la tabla de Detalle de Autores.
- Un Tipo de Perfil puede estar contenido en varias tuplas de la tabla de Detalle de URL
- Un Tipo de Perfil puede estar contenido en varias tuplas de la tabla de Detalle de Temas

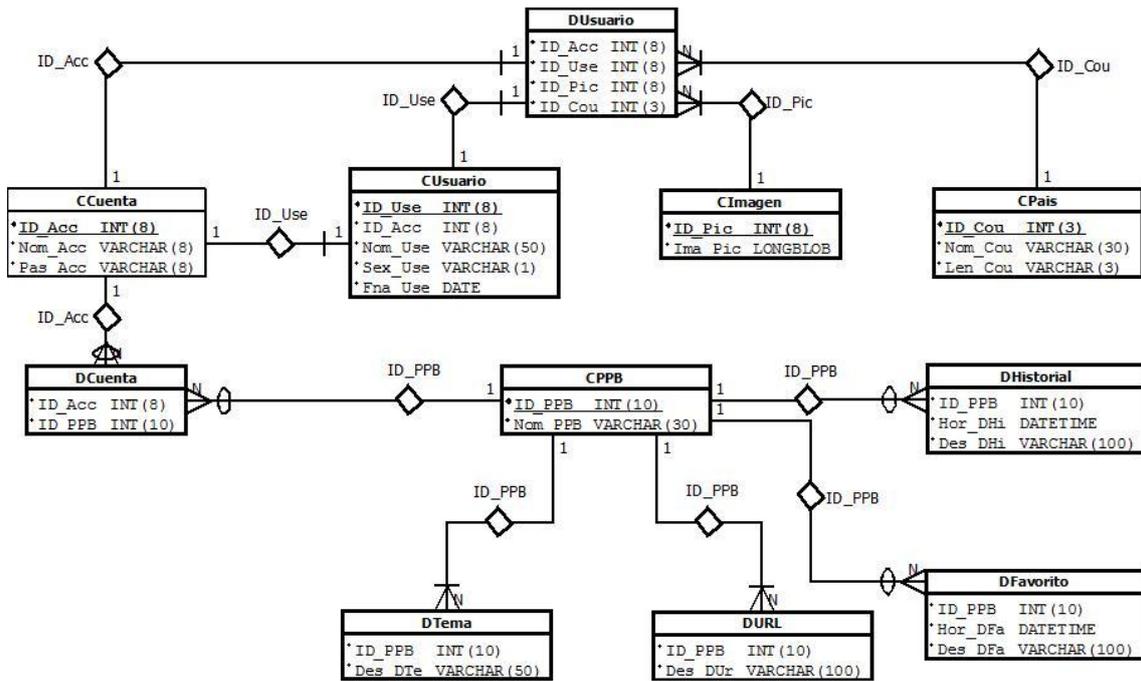


Figura 24: Diagrama Relacional

## 5.9 Pantallas

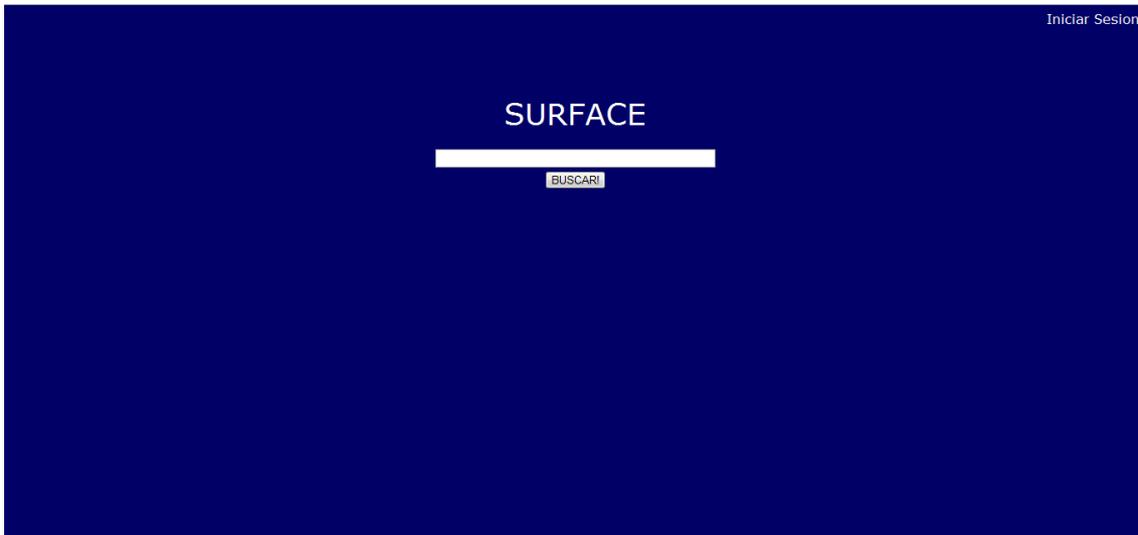


Figura 25: Pantalla de Búsqueda



Figura 26: Pantalla de Búsqueda Usuario Registrado

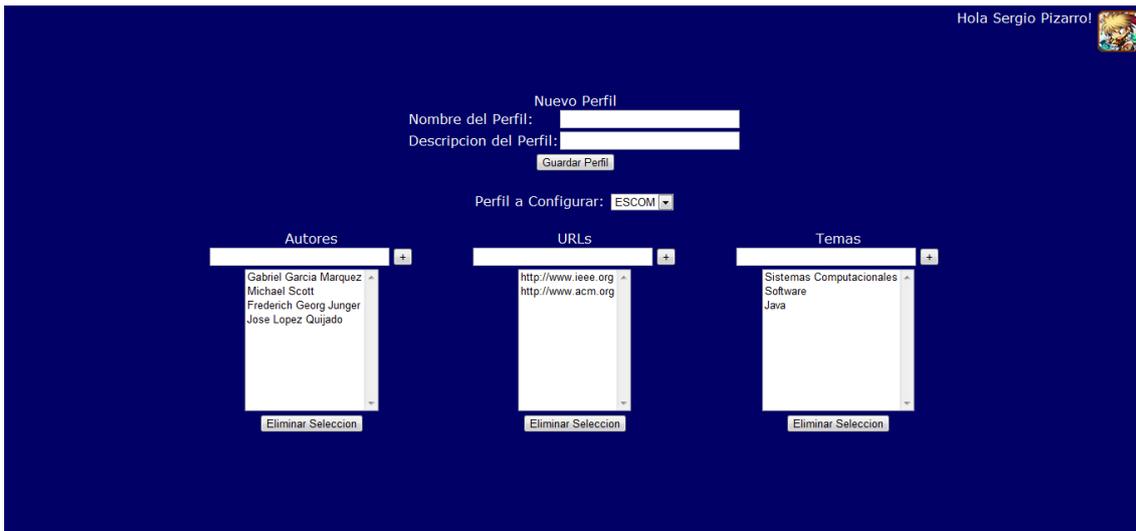


Figura 27: Pantalla de Registro de Perfil

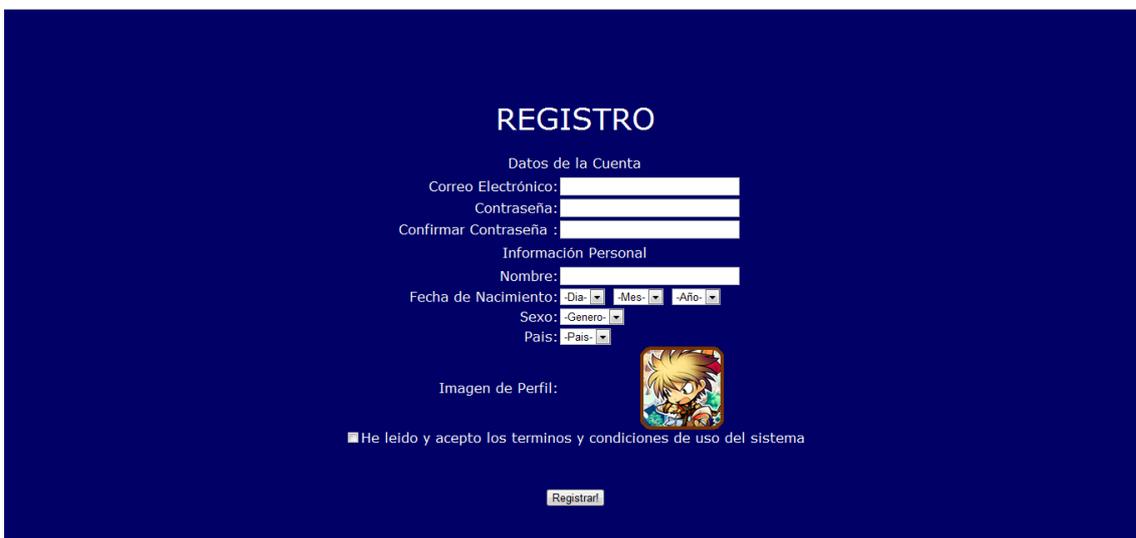


Figura 28: Pantalla de Registro de Usuario



Figura 29: Pantalla de Resultados de Búsqueda

## Referencias

### Citas

[1].-<http://www.abc.es/20110214/medios-redes/abci-bing-busquedas-precisas-201102141247.html>

[2].-<http://www.redalyc.uaemex.mx/pdf/114/11400508.pdf>

[3].- <http://www.redalyc.uaemex.mx/pdf/114/11400508.pdf>

[4].-<http://www.bloginformatica.net/el-porcentaje-de-uso-de-internet-explorer-esta-por-primera-vez-en-menos-del-60/3009/>

### Antecedentes

#### Fuentes de marco histórico

- Navegadores
  - [http://es.wikipedia.org/wiki/Navegador\\_web](http://es.wikipedia.org/wiki/Navegador_web)
  - <http://www.masadelante.com/faqs/que-es-un-navegador>
  
- Servidores
  - [http://es.wikipedia.org/wiki/Servidor\\_web](http://es.wikipedia.org/wiki/Servidor_web)
  - <http://www.pergaminovirtual.com/definicion/Servidor.html>
  - <http://www.editum.org/Que-Es-Un-Servidor-Web-p-401.html>
  - <http://www.masadelante.com/faqs/servidor>
  - <http://www.josecriado.com/hosting-alojamiento-web/caracteristicas-de-un-servidor-web/>
  
- Host
  - [http://es.wikipedia.org/wiki/Alojamiento\\_web](http://es.wikipedia.org/wiki/Alojamiento_web)
  - <http://www.masadelante.com/faqs/host>
  - <http://www.masadelante.com/faqs/hosting>

## **Análisis del sistema**

### **Fuentes de arquitectura de Software**

- [http://es.wikipedia.org/wiki/Arquitectura\\_de\\_software](http://es.wikipedia.org/wiki/Arquitectura_de_software)

### **Fuentes de modelo de arquitectura que se usara**

#### Modelos Existentes

- [ftp://190.5.199.3/jjurado/Ingenieria%20de%20Software%20II/Arquitectura\\_Software.pdf](ftp://190.5.199.3/jjurado/Ingenieria%20de%20Software%20II/Arquitectura_Software.pdf)

#### Modelo 3 Capas

- <http://www.managinf.com/arquitectura.pdf>
- [http://www.docirs.cl/arquitectura\\_tres\\_capas.htm](http://www.docirs.cl/arquitectura_tres_capas.htm)

### **Fuentes de estilo de arquitectura**

#### Estilos Existentes

- <http://homepage.mac.com/imaz/iblog/C612772037/E20050907222635/Media/Algunos%20Tipos%20de%20Arquitecturas.pdf>
- [ftp://190.5.199.3/jjurado/Ingenieria%20de%20Software%20II/Arquitectura\\_Software.pdf](ftp://190.5.199.3/jjurado/Ingenieria%20de%20Software%20II/Arquitectura_Software.pdf)

#### Estilo Llamada y Respuesta (call and return)

- [ftp://190.5.199.3/jjurado/Ingenieria%20de%20Software%20II/Arquitectura\\_Software.pdf](ftp://190.5.199.3/jjurado/Ingenieria%20de%20Software%20II/Arquitectura_Software.pdf)
- <http://homepage.mac.com/imaz/iblog/C612772037/E20050907222635/Media/Algunos%20Tipos%20de%20Arquitecturas.pdf>

## **Lecturas Recomendadas**

Existen numerosas referencias y publicaciones donde se puede obtener gran cantidad de información relacionada. A nivel general recomendamos las lecturas del artículo [Wooldridge 1995], el cual puede considerarse como un clásico dentro de la temática, y también el primer capítulo del libro de [Huhns 1998] y los artículos de [Nwana 1996] y [Franklin 1996].

En cuanto a referencias más específicas sobre agentes de información podemos sugerir los libros publicados por [Caglayan 1997] y [Klusch 1999] donde se recogen los últimos avances sobre el tema.

Finalmente, para disponer de más información referente a aplicaciones existentes hoy en día, pueden ser consultadas las siguientes páginas Web:

- BargainFinder. URL: [http:// bf.cstar.ac.com/](http://bf.cstar.ac.com/)
- Bullseye. URL: <http://www.intelliseek.com/prod/bullseye.htm>
- CiteSeer. URL: <http://www.researchindex.com>
- Kasbah - Market Maker. URL: <http://ecommerce.media.mit.edu/maker/maker.htm>
- Letizia. URL: <http://lcs.www.media.mit.edu/people/lieber/Lieberary/Letizia/Letizia-Intro.html>

## Referencias de Autores

- [Caglayan 1997] Caglayan, Alper y Harrison, Colin: Agent sourcebook. New York, etc., Wiley Computer Publishing, cop. 1997.
- [Klusch 1999] Klusch, Matthias: Intelligent Information Agents. Springer-Verlag. 1999