

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
SANTIAGO - CHILE



“DISEÑO E IMPLEMENTACIÓN DE UN MOTOR DE BÚSQUEDA SOBRE LA RED CHILENA ACTUAL”

FRANCISCO IGNACIO PÉREZ CASTRO

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN INFORMÁTICA

Profesor Guía: Marcelo Mendoza
Profesor Correferente: José Luis Martí

Septiembre - 2019

DEDICATORIA

Primero que nada, quiero dedicar esta memoria y todo el trabajo realizado a todas aquellas personas que contribuyeron de una u otra forma en mi formación como Ingeniero.

Ya sea en la universidad, en el estudio en casa, en mis estudios de intercambio, en las actividades de recreación y en todas las experiencias que pasé para poder llegar a este momento. Me siento demasiado agradecido de Dios y de la vida por haber tenido la oportunidad de estudiar la carrera que me gusta en la casa de estudios que me gusta.

AGRADECIMIENTOS

En esta sección quisiera agradecer a todas aquellas personas que han sido un pilar fundamental en mi vida universitaria y que siempre me han dado el apoyo para seguir adelante con todas las metas que me he propuesto en la vida.

Para comenzar, me gustaría agradecer a cada una de las personas que componen mi querida familia. Mis padres Victor y Mirna siempre me han dado un apoyo y un cariño incondicional, en los buenos y malos momentos, gracias a ellos nunca me faltó algo. Agradecer también a mi hermana Valentina, que siempre me ha dado un apoyo y confianza incondicional, a tal punto de considerarla una de mis mejores amigas. A mi tío Francisco, mi tía Roxana y a mi tía abuela Miguelina que también siempre me han apoyado incansablemente en todas las etapas de mi vida. A mi primo Julio por el cariño y la inspiración que siempre me dio, siempre lo consideré como mi hermano mayor.

También quiero agradecer a mis amigos más cercanos, Dwight, Constanza, Nicolas, Jean-Baptiste, Sofia, Francisco, Fabian, Alfredo, José, Sebastian, Giuseppe, Francisca y a todas esas personas que hicieron de mi vida universitaria la mejor de todas. Todas esas personas ya las considero parte de mi familia.

Agradecer a todos los profesores que contribuyeron de una u otra forma a mi formación profesional y siempre estuvieron dispuestos a ayudarme.

Por último, agradecer a Dios, por siempre darme la fuerza y las mejores oportunidades para surgir en la vida, siempre estaré eternamente agradecido por todo lo que me ha dado.

RESUMEN

Resumen— Este trabajo de título tiene como objetivo diseñar e implementar un motor de búsqueda sobre la red chilena que existe actualmente y usando las mejores tecnologías disponibles hasta la fecha.

Para lograr esto se diseña un sistema capaz de rastrear sitios web chilenos de una manera eficiente, que permite construir una colección representativa del país, que luego es indexada y usada para responder a las consultas ingresadas por los usuarios en el buscador.

Luego, esta implementación se disponibiliza a través de un servidor web llamado *Inquiro*. Los resultados más significativos son el nivel de representación de los sitios chilenos, ya que prácticamente toda la colección corresponde a sitios de Chile. También se obtiene que el algoritmo de *ranking* obtiene un nivel de precisión y desempeño competitivo con respecto a buscadores similares. Como resultado final, este sistema también permitirá la ejecución de experimentos e investigaciones futuras de alto impacto.

Palabras Clave— Motor de búsqueda, Dominios chilenos, Rastreador, Ranking, Documentos.

ABSTRACT

Abstract— This title work aims to design and implement a search engine on the Chilean network that currently exists and using the best technologies available to date.

To achieve this, a system capable of tracking Chilean websites in an efficient manner is designed, which allows building a representative collection of the country, which is then indexed and used to respond to queries entered by users in the search engine. Then, this implementation is made available through a web server called *Inquiro*.

The most significant results are the level of representation of Chilean sites, since practically the entire collection corresponds to sites in Chile. It is also obtained that the ranking algorithm obtains a level of accuracy and competitive performance with respect to similar search engines. As a final result, this system will also allow the execution of future high impact experiments and research.

Keywords— Search Engine, Chilean Domains, Crawler, Ranking, Documents.

GLOSARIO

DI: Departamento de Informática.

UTFSM: Universidad Técnica Federico Santa María.

API: Application Programming Interface. Conjunto de funciones y procedimientos que cumplen una o muchas funciones con el fin de ser utilizadas por otro software.

IR: Information Retrieval. Ciencia que estudia la búsqueda y recuperación de información en documentos electrónicos y cualquier tipo de colección de documentos digitales.

REST: Representational State Transfer. Estilo de arquitectura que comunica sistemas computacionales en la web. Se caracterizan por separar la lógica correspondiente al cliente y al servidor.

CSV: Comma-Separated Values. Tipo de documento con un formato sencillo para representar datos en forma de tabla, en las que las columnas se separan por comas.

XML: EXtensible Markup Language. Meta-lenguaje que permite definir lenguajes de marcas desarrollado por el World Wide Web Consortium (W3C) utilizado para almacenar datos en forma legible.

JSON: JavaScript Object Notation. Formato de texto sencillo para el intercambio de datos.

WWW: World Wide Web. Sistema lógico de acceso y búsqueda que utilizan las páginas web.

HTTP: Hypertext Transfer Protocol. Protocolo usado para visualizar paginas web.

FTP: File Transfer Protocol. Protocolo usado para transferir archivos.

JVM: Java Virtual Machine. Máquina virtual que permite que una computadora ejecute programas Java, así como programas escritos en otros lenguajes que también están compilados en el código de bytes de Java.

IDE: Entorno de Desarrollo Integrado. Aplicación informática que proporciona servicios integrales para facilitarle al desarrollador o programador el desarrollo de software.

ÍNDICE DE CONTENIDOS

| | |
|--|----|
| RESUMEN | IV |
| ABSTRACT | IV |
| GLOSARIO | V |
| ÍNDICE DE FIGURAS | IX |
| ÍNDICE DE TABLAS | IX |
| INTRODUCCIÓN | 1 |
| CAPÍTULO 1: DEFINICIÓN DEL PROBLEMA | 2 |
| 1.1 Situación Actual | 2 |
| 1.2 Árbol del Problema | 2 |
| 1.2.1 Amplio interés en búsquedas generales | 2 |
| 1.2.2 Búsquedas poco variadas de contenido chileno | 3 |
| 1.3 Motor de Búsqueda Inquirol | 4 |
| 1.4 Objetivos | 5 |
| 1.4.1 Objetivo General | 5 |
| 1.4.2 Objetivos Específicos | 5 |
| CAPÍTULO 2: MARCO CONCEPTUAL | 6 |
| 2.1 Recuperación de la información | 6 |
| 2.2 Tipos de buscadores | 7 |
| 2.3 Componentes de un motor de búsqueda | 7 |
| 2.3.1 Crawler | 7 |
| 2.3.2 Parseador | 8 |
| 2.3.3 Índice | 10 |
| 2.3.4 Buscador | 11 |
| 2.4 Ranking | 13 |
| 2.4.1 Modelo Booleano | 14 |
| 2.4.2 Modelo de Vectores Espaciales | 14 |
| 2.4.3 Modelo de Lenguaje Estadístico | 17 |
| CAPÍTULO 3: PROPUESTA DE SOLUCIÓN | 19 |
| 3.1 Diseño de la Solución | 19 |
| 3.2 Elección de Tecnologías | 20 |
| 3.2.1 Crawler | 20 |
| 3.2.2 Motor de Búsqueda | 23 |
| 3.2.3 Framework Interfaz Gráfica | 25 |

| | | |
|-------------------------|--|-----------|
| 3.3 | Diseño General de Componentes | 25 |
| 3.3.1 | Estructura General del sistema | 26 |
| 3.3.2 | Campos de extracción escogidos | 30 |
| 3.4 | Diseño de Crawler.py | 31 |
| 3.4.1 | Scrapy | 31 |
| 3.4.2 | Spider | 33 |
| 3.4.3 | HTML Parser | 36 |
| 3.4.4 | File Parser | 39 |
| 3.4.5 | Tika | 40 |
| 3.5 | Diseño de Main.cpp | 41 |
| 3.5.1 | Controlador | 42 |
| 3.5.2 | Crawler.py | 49 |
| 3.5.3 | Solr | 50 |
| 3.5.4 | Spdlog | 54 |
| 3.6 | Diseño de Interfaz | 55 |
| 3.6.1 | Vista principal de Escritorio | 55 |
| 3.6.2 | Vista principal Móvil | 56 |
| 3.6.3 | Vista de Resultados de Escritorio | 56 |
| 3.6.4 | Vista de Resultados Móvil | 57 |
| 3.7 | Implementación | 58 |
| 3.7.1 | Recursos utilizados | 58 |
| 3.7.2 | Implementación General | 60 |
| CAPÍTULO 4: | VALIDACIÓN DE LA SOLUCIÓN | 63 |
| 4.1 | Comparación de Inquiro con Buscadores Actuales | 63 |
| 4.1.1 | Comparación de consultas informacionales | 65 |
| 4.1.2 | Comparación de consultas navegacionales | 67 |
| 4.1.3 | Comparación de consultas transaccionales | 69 |
| 4.1.4 | Comparación cualitativa de Resultados | 71 |
| 4.2 | Estadísticas de la Colección | 75 |
| 4.2.1 | Tiempo de Recolección | 76 |
| 4.2.2 | Tamaño de la Colección | 76 |
| 4.2.3 | Cantidad de sitios según tipo de documento | 76 |
| 4.2.4 | Profundidad Promedio | 77 |
| 4.2.5 | Estadísticas con herramienta Luke | 78 |
| CAPÍTULO 5: | CONCLUSIONES | 84 |
| ANEXOS | | 89 |
| 6 ANEXO A: | RESULTADOS DE CONSULTAS BUSCADORES | 89 |
| 6.1 | Resultados Inquiro | 89 |
| 6.1.1 | Consulta: profesores usm | 89 |
| 6.1.2 | Consulta: eclipse solar 2019 | 91 |

| | | |
|---|--|------------|
| 6.1.3 | Consulta: radio carolina | 94 |
| 6.1.4 | Consulta: red transantiago | 96 |
| 6.1.5 | Consulta: comprar caliente camas | 99 |
| 6.1.6 | Consulta: 13c vivo online | 101 |
| 6.2 | Resultados Google | 103 |
| 6.2.1 | Consulta: profesores usm | 103 |
| 6.2.2 | Consulta: eclipse solar 2019 | 105 |
| 6.2.3 | Consulta: radio carolina | 108 |
| 6.2.4 | Consulta: red transantiago | 109 |
| 6.2.5 | Consulta: comprar caliente camas | 111 |
| 6.2.6 | Consulta: 13c vivo online | 113 |
| 6.3 | Resultados Yahoo | 116 |
| 6.3.1 | Consulta: profesores usm | 116 |
| 6.3.2 | Consulta: eclipse solar 2019 | 119 |
| 6.3.3 | Consulta: radio carolina | 121 |
| 6.3.4 | Consulta: red transantiago | 124 |
| 6.3.5 | Consulta: comprar caliente camas | 127 |
| 6.3.6 | Consulta: 13c vivo online | 129 |
| 6.4 | Resultados Bing | 132 |
| 6.4.1 | Consulta: profesores usm | 132 |
| 6.4.2 | Consulta: eclipse solar 2019 | 135 |
| 6.4.3 | Consulta: radio carolina | 138 |
| 6.4.4 | Consulta: red transantiago | 140 |
| 6.4.5 | Consulta: comprar caliente camas | 143 |
| 6.4.6 | Consulta: 13c vivo online | 146 |
| REFERENCIAS BIBLIOGRÁFICAS | | 150 |

ÍNDICE DE FIGURAS

| | | |
|----|---|----|
| 1 | Árbol del problema de la memoria. | 3 |
| 2 | Arquitectura de un sistema IR general. | 6 |
| 3 | Árbol de parseo DOM para una página web simple. | 9 |
| 4 | Diagrama Arquitectónico de <i>Inquiro.cl</i> | 20 |
| 5 | Diagrama de Componentes de script <i>Crawler.py</i> | 31 |
| 6 | Estructura estándar de documentos HTML | 37 |
| 7 | Diagrama de Componentes de script <i>Main.cpp</i> | 41 |
| 8 | Mensaje desplegado por el comando <i>help</i> del <i>Makefile</i> | 43 |
| 9 | Mensaje desplegado por <i>main.cpp</i> al ejecutar el <i>crawler</i> | 44 |
| 10 | Mensaje desplegado por <i>main.cpp</i> luego de terminar una acción. | 44 |
| 11 | Mensaje desplegado por <i>main.cpp</i> luego de reiniciar la colección de <i>SolR</i> | 46 |
| 12 | Mensaje desplegado por <i>main.cpp</i> al importar archivos <i>CSV</i> | 47 |
| 13 | Interfaz gráfica de la páginas web principal de <i>Inquiro.cl</i> | 56 |
| 14 | Interfaz gráfica de la páginas web móvil principal de <i>Inquiro.cl</i> | 56 |
| 15 | Interfaz gráfica de la páginas web de Resultados de <i>Inquiro.cl</i> | 57 |
| 16 | Interfaz gráfica de la páginas web móvil de Resultados de <i>Inquiro.cl</i> | 57 |
| 17 | Comparación del N° de resultados de dos consultas informacionales | 66 |
| 18 | Comparación del tiempo de respuesta de dos consultas informacionales | 67 |
| 19 | Comparación del N° de resultados de dos consultas navegacionales | 67 |
| 20 | Comparación del tiempo de respuesta de dos consultas navegacionales | 68 |
| 21 | Comparación del N° de resultados de dos consultas transaccionales | 69 |
| 22 | Comparación del tiempo de respuesta de dos consultas transaccionales | 70 |

ÍNDICE DE TABLAS

| | | |
|----|--|----|
| 1 | Dos índices invertidos, el primero es más simple, el segundo es una versión más compleja | 11 |
| 2 | Comparación de características de crawlers populares. | 21 |
| 3 | Comparación de desempeño práctico de crawlers populares. | 21 |
| 4 | Comandos Makefile para ejecutar en el Controlador. | 42 |
| 5 | Campos de la colección de SolR. | 51 |
| 6 | Recursos de Software utilizados en el desarrollo de <i>Inquiro.cl</i> | 59 |
| 7 | Recursos de Hardware de Inquiro. | 60 |
| 8 | Consultas aleatorias diseñadas para la comparación de los buscadores. | 64 |
| 9 | Comparación de resultados en distintos buscadores. | 65 |
| 10 | Consultas aleatorias diseñadas para la comparación de los buscadores. | 77 |
| 11 | Los 50 términos más frecuentes del campo <i>Cuerpo (Body)</i> | 80 |
| 12 | Los 50 términos más frecuentes del campo <i>Título (Title)</i> | 81 |
| 13 | Los 50 términos más frecuentes del campo <i>Texto Ancla (Anchor text)</i> | 82 |
| 14 | Los 50 términos más frecuentes del campo <i>Dominio (Domain)</i> | 83 |

INTRODUCCIÓN

Hoy en día la necesidad de información se hace cada vez más evidente, ya que cada vez más, la llamada transformación digital, cobra más fuerza en el mundo y en Chile. Internet ha tenido un crecimiento explosivo a lo largo de los años y se puede decir que llega a tener del orden de los cientos de millones de documentos [Netcraft, 2019]. Es por esto que la velocidad del flujo del conocimiento debe ser más rápida, y los motores de búsqueda web juegan un papel muy importante en este sentido, ya que son los que se encargan de recorrer la web y guardar los enlaces a cada documento, de forma que si un usuario realiza una consulta que se relacione con estos documentos, estos salgan listados entre las opciones de interés.

Para que un buscador funcione como corresponde, deben existir componentes que recorren la internet, comúnmente llamados *crawlers*, y además debe poseer algoritmos de recuperación de documentos relevantes para cada consulta ingresada.

Según [Liu, 2007], la estructura principal de todo motor de búsqueda es la siguiente.

- Crawler
- Indexador
- Buscador
- Ranking

Estructura que se explicará en detalle en los capítulos siguientes.

Actualmente no existen tecnologías de búsqueda focalizadas sobre la web chilena , es decir, no existe una búsqueda de contenido centrado únicamente en los dominios “.cl”, lo que es un problema para ciertos grupos de la población que buscan contenidos nacionales más específicos.

En el pasado, existió un buscador llamado *Inquirol.CL* que fue desarrollado en los tiempos en que los motores de búsqueda populares estaban recién partiendo. Esto se usa como inspiración para el diseño de un motor de búsqueda vertical que intenta solucionar el problema de fondo. La implementación de este buscador se logra a través de la extracción de información de páginas web nacionales, con ayuda de un *crawler* diseñado específicamente para este propósito, y que permite la construcción de una gran colección de documentos, que luego son indexados eficientemente para una recuperación rápida de sitios relevantes.

Este trabajo se relaciona fuertemente con el área informática llamada Recuperación de Información, que estudia en detalle la manera en que se realizan las búsquedas de colecciones digitales. Por lo mismo, el usar tecnologías actuales en la implementación de este buscador, permite que la documentación de esta memoria sea de gran ayuda para futuras investigaciones que necesiten experimentar en esta área.

CAPÍTULO 1

DEFINICIÓN DEL PROBLEMA

1.1. Situación Actual

El mundo moderno en el que nos encontramos actualmente ya está muy acostumbrado a la tecnología, a las cosas inmediatas, noticias instantáneas, en donde la rapidez de información fluye cada vez con mayor facilidad. También esto a veces se puede volver un problema, las personas se vuelven cada vez más dependientes de la tecnología, ya sea por mera diversión o por negocios, se frustran si se produce lentitud en sus acciones o en sus búsquedas.

Si se analiza esta necesidad enfocándose en la realidad que existe hoy en día en la web chilena, se puede visualizar que las búsquedas de información usuales son siempre muy generales, es decir, se centran en lo que la mayoría de la gente de Chile busca en internet, pero no necesariamente en el contenido chileno en sí.

De esto trata el problema, de que actualmente no existen tecnologías de búsqueda focalizadas sobre la web chilena o, en otras palabras, no existe una búsqueda de contenido centrado únicamente en los dominios “.cl”. Esto es un problema porque una parte importante de la población de este país necesita una búsqueda de contenidos más específica y el uso de las tecnologías actuales no restringen los resultados de las búsquedas a este dominio, sino que a lo que busca un chileno en general.

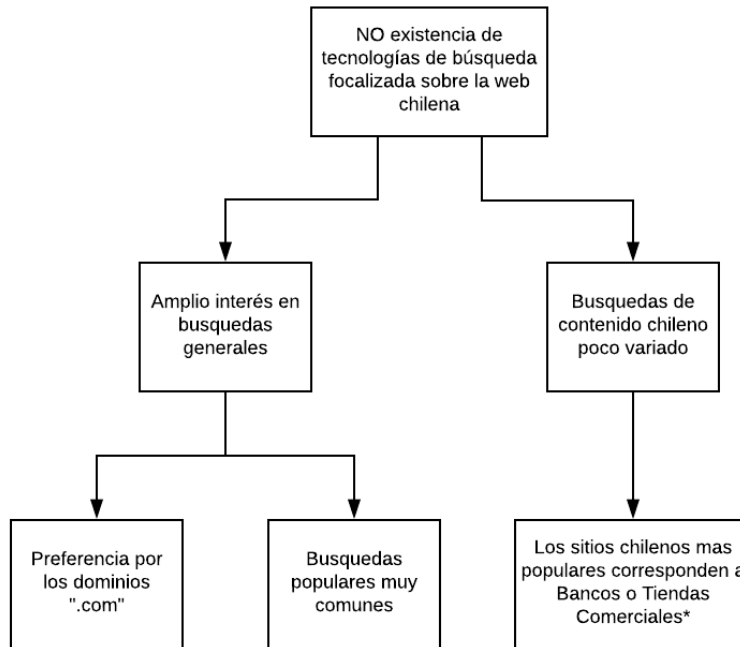
1.2. Árbol del Problema

Para visualizar mejor el problema y sus causas, se diseña un “árbol del problema” y se presenta en la figura 1.

Como se puede apreciar, existen dos causas principales de que no existan tecnologías de búsqueda chilenas. La primera es que exista mucho interés de parte de la población del país en búsquedas generales y la segunda es que las búsquedas de contenido chileno es poco variado. A continuación se procede a detallar estas causas.

1.2.1. Amplio interés en búsquedas generales

Actualmente las personas ven a los buscadores como una herramienta para encontrar la información que necesitan de una forma rápida y eficaz. Y la mayoría de las veces esta necesidad proviene de fuentes no exclusivamente chilenas, es decir, la mayoría de la información que un chileno promedio necesita proviene de dominios “.com”.



*: <https://www.alexa.com/topsites/countries/CL>

Figura 1: Árbol del problema de la memoria.

Fuente: Elaboración Propia.

Esto no necesariamente representa un problema, sino más bien una realidad: Los buscadores actualmente entregan resultados de lo que busca un chileno en general, no dándole tanta importancia al dominio de donde venga la información.

Lo que evidencia una segunda subcausa: las búsquedas populares chilenas son muy comunes, es decir, un chileno promedio tiende a buscar cosas similares a los de sus compatriotas. Lo cual no es extraño, ya que este fenómeno está presente en la mayor parte de los países del mundo.

1.2.2. Búsquedas poco variadas de contenido chileno

Esta causa se relaciona bastante con la de que exista un amplio interés con las búsquedas generales. Y es que según la página web Alexa ([Alexa, 2019]), sitio web de Amazon¹, hasta la fecha la mayoría de las personas que utilizan las páginas o contenido chileno solo buscan tiendas comerciales o bancos exclusivos de Chile, es decir, las personas están obligadas a acudir a sus sitios web para gestionar sus compras de la vida diaria.

¹<https://www.amazon.com/>

1.3. Motor de Búsqueda Inquiro.CL

Esta memoria, basará el problema y su solución en la implementación de un antiguo motor de búsqueda antes implementado en Chile y que solucionaba eficazmente el problema descrito. Se trata de la publicación [Mendoza *et al.*, 2009], motor de búsqueda que marcó un antes y un después en el tema de los motores de búsqueda en Chile, ya que se presenta un nuevo buscador para el país: *Inquiro.CL*. Este nuevo motor de búsqueda permitía a los usuarios buscar en la web latinoamericana (específicamente, en diez países de habla hispana) especificando su dominio de búsqueda (".cl", ".com.ar", ".com.mx" y el resto de América Latina), pero que hoy en día no está en funcionamiento.

Muchas características de este trabajo son destacables y sirven como base para proyectos como esta memoria. Por ejemplo, la estructura de este buscador se basaba en una arquitectura distribuida que gestiona dos colecciones de documentos. El primero era un caché de páginas o sitios que proporciona respuestas a bajo costo computacional. El segundo es una colección general de documentos que se visita cuando el usuario extiende la búsqueda a la segunda página de resultados o más allá.

El índice de búsqueda se creó utilizando una estrategia de múltiples niveles, de modo que los títulos, las URL, los encabezados y los contenidos completos del sitio que conforman la colección se catalogan en índices diferentes.

El motor de búsqueda utilizaba una función de clasificación que combina varias medidas de relevancia, entre ellas las preferencias del usuario, la clasificación de la página y el texto. En ese tiempo, la colección principal de *Inquiro* alcanzaba más de 1.500.000 páginas y aproximadamente 35.000 sitios. Además, los resultados experimentales mostraron que el motor de búsqueda era preciso y se comparaba favorablemente con los motores de búsqueda similares, alcanzando un promedio de más del 60 % de precisión entre los 5 mejores buscadores que existían a la fecha.

Es por esto que el desafío hoy en día es imitar el trabajo bien hecho y las buenas prácticas implementadas en *Inquiro* pero con la tecnología que existe actualmente.

1.4. Objetivos

1.4.1. Objetivo General

El objetivo general de esta memoria es diseñar e implementar un motor de búsqueda vertical para la búsqueda de contenidos específicos en la web chilena.

1.4.2. Objetivos Específicos

- Consolidar una colección representativa del dominio “.cl”.
- Embeber la colección consolidada en una plataforma de búsqueda.
- Diseñar una interfaz gráfica para el motor de búsqueda.
- Realizar pruebas de validación para analizar los resultados de búsqueda obtenidos.
- Disponibilizar el motor de búsqueda a través de un servidor en línea.

CAPÍTULO 2

MARCO CONCEPTUAL

En este capítulo se explicarán los conceptos y tecnologías que permiten desarrollar una solución efectiva para el problema antes planteado.

2.1. Recuperación de la información

El primer concepto importante en la explicación de esta memoria es el de la **Recuperación de la información (IR)**, es la base de todo motor de búsqueda, y según [Liu, 2007] se define como el estudio que ayuda a los usuarios a encontrar información que satisface sus necesidades de información. En otras palabras, IR estudia la adquisición, organización, el almacenamiento, la recuperación y la distribución de la información.

A lo largo del tiempo IR se ha enfocado principalmente a la recuperación de documentos, usando estos como unidad básica. En la figura 2) se puede visualizar el esquema general de un sistema IR.

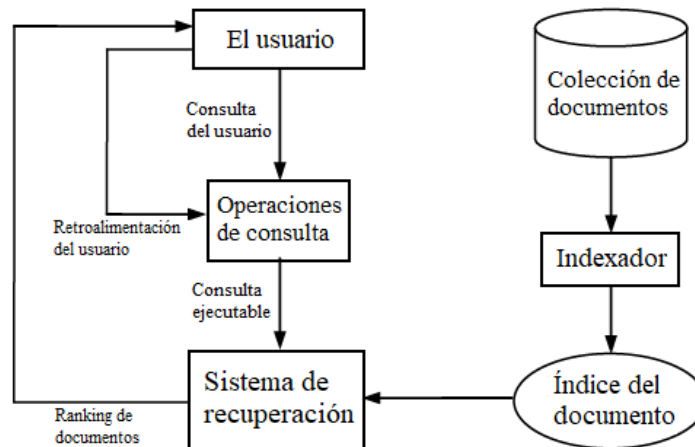


Figura 2: Arquitectura de un sistema IR general.
Fuente: [Liu, 2007].

Básicamente, el flujo es el siguiente, el usuario con necesidades de información realiza una consulta (Consulta del usuario) al Sistema de Recuperación a través del Módulo de Operaciones de Consulta. Este módulo a la vez usa el Índice del documento para recuperar aquellos documentos que contienen alguno de los términos de la consulta, ya que estos podrían ser relevantes. Luego, se calculan puntajes de relevancia para ellos y entonces se realiza un ranking de los documentos más relevantes según estos puntajes. Finalmente este ranking de documentos es presentado al usuario.

2.2. Tipos de buscadores

A continuación, se procede a explicar el tipo particular de motores de búsqueda que se usarán en esta memoria. Según [Kowalska, 2016] aquellos buscadores que realizan búsquedas generales se conocen como **buscadores horizontales**. Se dice que un buscador o motor de búsqueda es horizontal, si este recorre toda la web en busca de información general que cubra una amplia gama de temas. Pueden indexar grandes porciones de Internet utilizando un *crawler* web general. Como resultado, reúnen respuestas coincidentes para la mayor cantidad de consultas posibles. Ejemplos de buscadores horizontales son justamente los más conocidos, como lo son Google, Yahoo y Bing.

Por otro lado, los **buscadores verticales** o también llamados búsquedas por tópicos, están especializados a un tipo de medio, actualidad, zona geográfica o género de contenido. Significa que solo buscan en una parte específica de la Internet, una porción vertical de la misma. Este tipo de motores de búsqueda utilizan *crawlers* especializados que intentan indexar solo páginas web relacionadas con un tema especial. Ejemplos de buscadores verticales corresponden a todos los que realizan búsquedas por secciones, como lo son Google Images, buscador de imágenes por excelencia; o Youtube, el buscador de vídeos más utilizado actualmente. Otro ejemplo de buscadores más especializados a la zona geográfica, o en este caso, a países es Yandex, el buscador más utilizado en Rusia; o Baidu, el buscador más utilizado en China.

Es por esto que la implementación de un motor de búsqueda vertical sería lo más adecuado para realizar búsquedas focalizadas en la red chilena.

2.3. Componentes de un motor de búsqueda

A continuación se presentan los componentes principales que debe tener todo motor de búsqueda para ser llamado como tal y cuales de estos son a los que hay que ponerle especial énfasis para lograr una solución eficaz.

2.3.1. Crawler

Los *crawlers*, arañas o rastreadores web, son programas que tienen como función descargar páginas web. Uno de los componentes más importantes de un motor de búsqueda, ya que estos deben almacenar el contenido de cada página web a ofrecer en los resultados de sus búsquedas. De manera similar a como las personas siguen *hyperlinks* desde una página a otra para navegar en la web, los rastreadores pueden seguir un conjunto enorme de páginas siguiendo una cierta cantidad de reglas, extrayendo al mismo tiempo todo el contenido deseado de cada página visitada.

La necesidad de estos rastreadores puede llegar a ser bastante alta dependiendo de lo que

se requiera, ya que pueden usarse para hacer análisis de documentos ya sea en la web como en repositorios *off-line*. Estos últimos documentos no suponen un gran problema para los análisis, puesto que generalmente estos no suelen cambiar, contrario a la situación de los documentos que se encuentran *on-line* que presentan una alta tasa de cambio, es decir, se añaden, se editan, se mueven y se eliminan páginas web constantemente, lo que aumenta la necesidad de rastreadores para sistemas que requieren una actualización constante, como lo son los motores de búsqueda.

Además, este componente puede ser usado en distintas áreas, una de las más utilizadas es la de "Inteligencia de Negocios". Muchas organizaciones suelen hacer análisis de sus potenciales colaboradores o incluso de sus competidores directos, obteniendo información de sus dominios y poder realizar seguimientos de interés. Otra aplicación de esta misma área es la de notificaciones, ya que los rastreadores pueden usarse para monitorear constantemente sitios que sean de interés para algún usuario o comunidad, de forma de que estas sean notificadas oportunamente cuando ocurra una actualización de información.

Como siempre, no todas las aplicaciones son de uso apropiado, ya que estos pueden usarse para fines malintencionados, como lo es la obtención de direcciones *e-mail* e información personal para el uso indiscriminado de *spammers* y ataques informáticos como el *phishing*.

Pero sin duda, la aplicación principal de los rastreadores es para el apoyo de los motores de búsqueda, ya que estos se alimentan principalmente de las páginas que se recolectan en la web y que son utilizadas para construir sus índices de búsqueda. Ejemplos de estos buscadores pueden ser los más conocidos, como lo son Google, Yahoo! o Bing, que utilizan rastreadores muy eficientes para reunir una alta cantidad de páginas independientemente de su contenido.

Los *crawlers* que se usarán en esta memoria no corresponden a rastreadores generales, sino más bien a los **crawlers preferenciales**, los que se caracterizan por centrarse en descargar paginas de cierto tipo o de ciertos dominios, es decir, son más focalizados.

En primer lugar un *crawler* comienza rastreando un conjunto de páginas iniciales o también llamados páginas semilla. Dentro de estas páginas se recopila todo el contenido deseado y se van obteniendo a través de esto mas y mas *hyperlinks* de páginas, siempre visitando y extrayendo información, así sucesivamente hasta que un número suficiente de páginas es alcanzado o se cumple alguna otra condición.

2.3.2. Parseador

Luego de que el *crawler* descarga las páginas rastreadas, viene el siguiente paso, que es parsear la información, es decir, extraer el contenido de las secciones específicas de cada documento, como lo son las etiquetas *HTML*, los títulos, subtítulos, cuerpo del documento, extracción de *hyperlinks*, entre otros. Todas estas secciones están descritas por el Modelo de

Objetos del Documento (DOM), que establece la estructura *HTML* de cada página web como un árbol de parseo y que se puede observar en la figura 3).

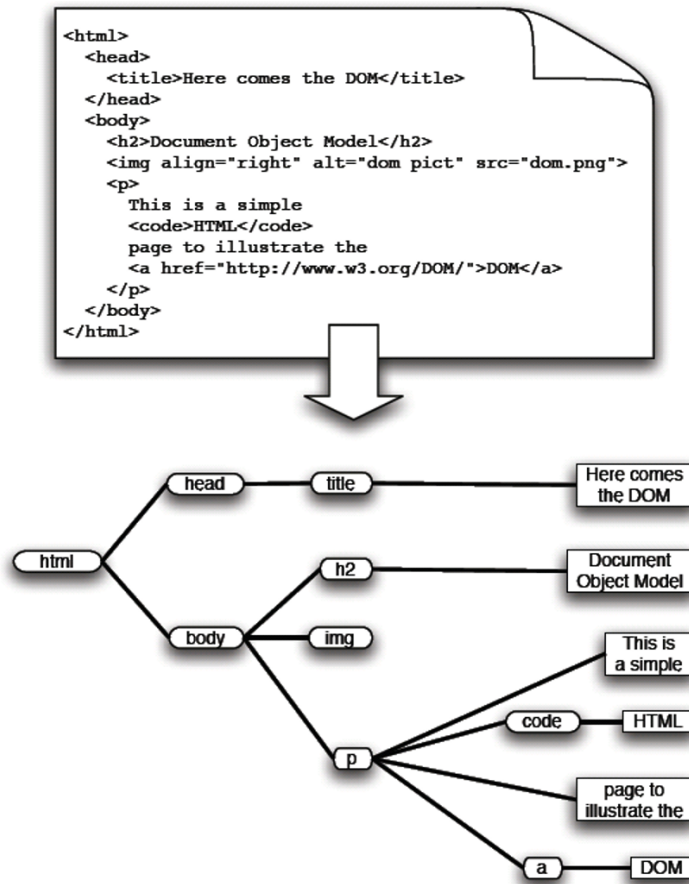


Figura 3: Árbol de parseo DOM para una página web simple.
Los nodos internos representan las etiquetas HTML y las hojas los valores correspondientes. Fuente: [Liu, 2007].

Sin embargo, no siempre es fácil parsear todo tipo de páginas web, ya que muchos autores olvidan poner ciertas etiquetas, las definen de manera incorrecta o incluso olvidan ciertos atributos requeridos. Es por esto que el proceso de parseo no debe ser muy riguroso con respecto a esto, ya que podrían dejarse afuera páginas importantes solo por esta razón.

Otro dato interesante es que los *crawlers* pueden parsear información en distintos formatos, tanto en *HTML* como en *XHTML*, pueden leer archivos planos, PDF, archivos *Microsoft Word* y *PowerPoint*, entre otros, lo que aumenta el número de información para recopilar considerablemente. Aunque a veces es incluso más complejo, como lo son los casos en que las páginas, por ejemplo, contienen animaciones *flash* o mapas que complican bastante la labor de estos rastreadores.

2.3.3. Índice

Este componente o etapa es la encargada de almacenar documentos en bruto en alguna estructura de datos eficiente para la posterior recuperación del contenido.

Existe un tipo particular de índice que se usa principalmente en los motores de búsqueda y en muchos sistemas IR, que se denomina el **índice invertido**. Este tipo de índice es famoso por su simplicidad y por ser muy eficiente para las búsquedas. Su funcionamiento consiste en que el sistema calcula un puntaje de relevancia para cada documento indexado con respecto a los términos ingresados en la consulta y son presentados al usuario ordenados por importancia.

La eficiencia de este índice también está dada por la forma en que se rankean los documentos, ya que no se compara cada término de la consulta con todos los documentos almacenados, si no que solo se analiza la relevancia de un pequeño conjunto de documentos que contienen al menos uno de los términos ingresados.

Más detalladamente, el índice invertido de una colección de documentos se compone de una estructura de datos que relaciona cada término en específico con una lista de todos los documentos que contienen dicho término, por lo tanto recuperar esta lista de documentos para un término en particular tomaría un tiempo lineal constante.

La búsqueda de documentos para múltiples términos en una consulta es bastante sencillo y se procede a explicar el proceso de una manera más formal. Dado un conjunto de documentos $D = \{d_1, d_2, \dots, d_N\}$ cada documento posee un identificador único (ID). Además, un índice invertido se compone de dos partes: un vocabulario V , que contiene todos los términos distintos en el conjunto de documentos, y por cada único término t_i se tiene una lista invertida de posteos. Cada posteo almacena el ID (j) del documento d_j que contiene el término t_i y cualquier otra información con respecto a t_i en d_j , es decir, dependiendo de las necesidades que debe retornar el algoritmo de ranking, se pueden añadir varios tipos de información. Por ejemplo, si se quiere centrar el algoritmo en la búsqueda por proximidad de cada frase ingresada, cada posteo para un término t_i sería de la forma:

$$\langle id_j, f_{ij}, [o_1, o_2, \dots, o_{|f_{ij}|}] \rangle \quad (1)$$

Donde id_j es el ID del documento d_j que contiene el término t_i , f_{ij} es un contador de frecuencia de t_i en d_j , y o_k son las posiciones del término t_i en d_j . Cada uno de estos posteos están ordenados en orden ascendente basado en las posiciones de estos en el documento.

Ejemplo 1: Tenemos tres documentos de id_1 , id_2 y id_3 :

ID 1 : La minería web es útil

ID 2 : El uso de aplicaciones de minería

ID 3 : La minería de la estructura de la web estudia la estructura de hiperenlace de la web

En el que cada término en el documento ocupa una posición respectiva, es decir, en el ejemplo, en el documento con ID 1, "minería" ocuparía la posición 2, "web" la posición 3 y así.

Mientras que el vocabulario sería el siguiente conjunto:

$\{minería, web, util, uso, aplicaciones, estructura, estudia, hiperenlace\}$

Al haber aplicado el proceso de parseo se realiza un preprocesamiento en el que se remueven, por ejemplo, las *stopwords* "es" y "la", palabras que no aportan valor a la búsqueda por carecer de significado.

| | | | |
|---------------|--------------------|--------------|--|
| Aplicaciones: | id_2 | Aplicaciones | $\langle id_2, 1, [4] \rangle$ |
| Hiperenlace: | id_3 | Hiperenlace | $\langle id_3, 1, [13] \rangle$ |
| Minería: | id_1, id_2, id_3 | Minería | $\langle id_1, 1, [2] \rangle, \langle id_2, 1, [6] \rangle, \langle id_3, 1, [2] \rangle$ |
| Estructura: | id_3 | Estructura | $\langle id_3, 2, [5, 11] \rangle$ |
| Estudia: | id_3 | Estudia | $\langle id_3, 1, [9] \rangle$ |
| Uso: | id_2 | Uso | $\langle id_2, 1, [2] \rangle$ |
| Útil: | id_1 | Útil | $\langle id_1, 1, [5] \rangle$ |
| Web: | id_1, id_3 | Web | $\langle id_1, 1, [3] \rangle, \langle id_3, 2, [8, 16] \rangle$ |
| (A) | | (B) | |

Tabla 1: Dos índices invertidos, el primero es más simple, el segundo es una versión más compleja

Fuente: [Liu, 2007].

Como se puede ver en la tabla 1, el índice invertido (A) representa una versión simple de este tipo de índice, el cual cada término está asociado simplemente con una lista de IDs de documentos que contienen el término. En cambio en el índice (B) existe más complejidad, puesto que contiene información adicional como un contador de frecuencia y sus posiciones en el documento.

2.3.4. Buscador

Las búsquedas que utilizan un índice invertido son muy particulares, ya que cada término de las consultas ingresadas es procesado para obtener de manera eficiente las listas invertidas respectivas y así encontrar aquellos documentos que contienen todos o muchos de los términos ingresados.

Más detalladamente, una vez que se ingresan términos en una consulta, se realiza lo que se denomina una "búsqueda por relevancia" y que consta de tres pasos principales:

1. Búsqueda de vocabulario

El vocabulario consta de todas las palabras, que no sean *stopwords*, y que son parte del conjunto de palabras indexadas en el índice invertido. Este paso se encarga de encontrar cada término de la consulta en este vocabulario para poder obtener las listas invertidas correspondientes. Usualmente, este vocabulario reside en memoria principal, para que de esta manera, se incremente la velocidad de procesamiento. Otra medida que aumenta la velocidad es la del uso de *hashing*, árboles o *B-trees* para el indexamiento, ya que estas estructuras de datos son muy utilizadas para realizar búsquedas eficientes. Un indexamiento ordenado lexicográficamente aumenta la eficiencia de espacio y además permite realizar búsquedas binarias, que poseen una complejidad $O(\log|V|)$, donde $|V|$ es el tamaño del vocabulario.

2. Combinación de resultados

Luego de que se obtienen las listas invertidas de cada término, la idea es intersectarlas para así encontrar la lista definitiva, es decir, el conjunto de documentos que contiene todos los términos de la consulta.

Existen algunas heurísticas importantes que permiten un mejor desempeño en la combinación, por ejemplo, comenzar con la lista más pequeña como base e ir combinando con las demás que son más largas. De esta manera, para cada posteo es la lista más corta se puede realizar una búsqueda binaria para encontrarlo en las listas más largas más rápidamente. También es posible entregar resultados parciales a la consulta, es decir, que contienen sólo algunos de los términos ingresados, lo que es más útil en la práctica, ya que no siempre lo que se busca contiene todas y cada una de las palabras ingresadas. Además, la mayoría de las veces todo el índice invertido no cabe en memoria principal, por lo que solo una parte de este es puesto en caché para incrementar la eficiencia.

Existe todo un estudio para decidir qué partes del índice invertido se deben poner en caché, y que involucran análisis de registros de consulta, para poder determinar términos de consulta frecuentes y poner las listas invertidas de estos términos en caché. Es importante recalcar que si la consulta realizada sólo contiene un solo término, la búsqueda pasa directamente al paso 3.

3. Cómputo del puntaje de ranking

El paso final de esta búsqueda es la de computar un puntaje de ranking de relevancia para cada uno de los documentos seleccionados en los pasos anteriores. El cómputo de estos puntajes se calculan en base a funciones de relevancia (por ejemplo, Coseno o Okapi), que podrían considerar la proximidad de la información en cuanto a los términos o la frase entera. Y que se profundiza más en detalle en la siguiente etapa.

2.4. Ranking

El algoritmo de ranking es una de las cosas más importantes que tiene que manejar un motor de búsqueda, ya que es lo que define que un buscador sea de calidad o no. Muchos de los algoritmos que se usan en los más grandes buscadores comerciales como Google, Yahoo! o Bing, son de difícil acceso y poca es la información que se tiene. Sin embargo, los algoritmos más utilizados se basan en una implementación preliminar de Google.

Los motores de búsqueda en general se basan en los modelos IR, ya que estos se encargan de controlar de qué manera un documento y una consulta son representadas y como es definida la relevancia de estos según las consultas ingresadas por los usuarios.

En cuanto a las búsquedas *web*, existen tres grandes modelos clásicos: El **modelo Booleano**, el **modelo de vectores espaciales** y el **modelo de lenguaje estadístico**. Si bien son distintos modelos, ya que representan los documentos y las consultas de diferente forma, todos comparten cosas en común, y es que usan el mismo *framework*, es decir, la misma manera de hacer las cosas.

Por ejemplo, todos estos modelos tratan cada documento o consulta como una "bolsa" de términos o palabras. La secuencia y posición en una oración o un documento son ignorados, es decir, un documento solo es descrito por un conjunto de términos distintivos. Un término en este contexto se puede considerar como una palabra cuya semántica ayuda a recordar el tema principal del documento. Además, en estos modelos cada término es asociado con un peso.

Dada una colección de documentos D , sea $V = \{t_1, t_2, \dots, t_{|V|}\}$ el conjunto de términos distintivos en la colección y que suele llamarse el **vocabulario** de la colección, donde t_i es un término y $|V|$ es el tamaño del vocabulario, es decir, el número de términos en V . Un peso $w_{ij} > 0$ es asociado con cada término t_i de un documento $d_j \in D$. Teniendo el valor $w_{ij} = 0$ para un término que no aparece en el documento d_j . Por lo tanto, cada documento d_j es representado con un vector de términos,

$$d_j = (w_{1j}, w_{2j}, \dots, w_{|V|j}) \quad (2)$$

donde cada peso w_{ij} corresponde al peso del término $t_i \in V$, que representa el nivel de importancia de t_i en el documento d_j .

Con esta representación, una colección de documentos es básicamente representado como una matriz o tabla relacional. Cada término es un atributo y cada peso es un valor de esos atributos. Lo que cambia en cada modelo de recuperación es la forma en que se computan los valores w_{ij} y que se explican a continuación.

2.4.1. Modelo Booleano

Este modelo es de los más simples que existen y es que también es de los más antiguos. Se basa en la noción de encontrar los documentos que contengan la consulta exacta ingresada. La recuperación y la consulta están basados en álgebra Booleana. La representación de los términos y de los documentos en este modelo corresponden a conjuntos de término, es decir, cada término sólo es considerado presente o ausente en un documento. El peso $w_{ij} (\in \{0, 1\})$ de un término t_i en un documento d_j es 1 si t_i aparece en d_j , y 0 en el caso contrario, es decir,

$$w_{ij} = \begin{cases} 1 & \text{si } t_i \text{ aparece en } d_j \\ 0 & \text{si no} \end{cases} \quad (3)$$

Consultas booleanas

Como se ha explicado anteriormente, los términos de la consulta se pueden combinar lógicamente usando operadores booleanos tales como AND, OR Y NOT, los cuales son los clásicos operadores con una semántica lógica y que implica a la vez que una consulta booleana tenga una semántica precisa. Por ejemplo, la expresión $((x \text{ AND } y) \text{ AND } (\text{NOT } z))$ implica que un documento recuperado debe obligatoriamente tener los términos x e y pero no el término z . Otro ejemplo sería la expresión $(x \text{ OR } y)$, que significa que al menos uno de esos términos debe estar en los documentos recuperados

Recuperación de documentos

Teniéndose una consulta booleana, el sistema recupera cada documento que hace a la consulta verdadera. Por lo tanto, todo se reduce a la decisión binaria de sí un documento es relevante o irrelevante, llamado también *coincidencia exacta*. Como su nombre lo dice, solo se toman en cuenta las coincidencias exactas, es decir, no se toman en cuenta las coincidencias cercanas o similares para la recuperación de esos documentos. Esta desventaja ocasiona que este modelo sea considerado sólo como un modelo de pruebas, ya que sus resultados son de mala calidad la mayoría de las veces.

Por otro lado, esto también implica que la frecuencia de términos y su proximidad contribuyen significativamente a la relevancia de un documento. A pesar de todo esto, los motores de búsqueda más populares actualmente soportan consultas booleanas limitadas a usar operadores para incluir y excluir ciertos términos de la consulta.

2.4.2. Modelo de Vectores Espaciales

Probablemente el modelo IR más usado y conocido. Posee una representación de documentos bien particular, ya que estos se representan como vectores de pesos, que se calculan basándose en una variación del esquema TF-IDF que explicaremos a continuación. No es como en el modelo booleano descrito anteriormente, ya que los pesos w_{ij} de los término t_i

en los documentos d_j ya no corresponden a los valores 0 o 1, sino que pueden ser cualquier número.

Esquema de Frecuencia de Términos (TF)

Este método representa cada peso de un término t_i en un documento d_j como el número de veces que t_i aparece en el documento d_j , denotado como f_{ij} .

Lo malo de este esquema es que no considera el caso en donde un término aparece en muchos documentos de la colección, y por lo tanto, podría no ser diferenciador.

Esquema de frecuencia inversa de documentos (TF-IDF)

Es el esquema más conocido y también se basa en la frecuencia de términos (TF). La diferencia está en que este esquema apuesta por la frecuencia inversa de documentos (IDF). Si bien existen muchas variaciones de este esquema, a continuación se explicará la versión más básica.

Si N es el número total de documentos en el sistema o colección, df_i el número de documentos en los que el término t_i aparece al menos una vez, y f_{ij} el contador de frecuencia bruta del término t_i en el documento d_j . Entonces la frecuencia de términos normalizada (tf_{ij}) de t_i en d_j está dada por:

$$tf_{ij} = \frac{f_{ij}}{\max \{f_{1j}, f_{2j}, \dots, f_{|V|j}\}} \quad (4)$$

En donde el denominador contiene una función máxima que se computa sobre todos los términos que aparecen en el documento d_j . Si el término t_i no aparece en d_j entonces $tf_{ij} = 0$. Y también recordar que $|V|$ es el tamaño del vocabulario de la colección.

Pero como se mencionó anteriormente, lo característico de este esquema es la frecuencia inversa de documentos de un término t_i , denotada como idf_i y que está dada por:

$$idf_i = \log \frac{N}{df_i} \quad (5)$$

Lo cual es lógico, puesto que si un término aparece en un gran número de documentos de la colección, probablemente no sea importante o poco discriminativa.

Y por último, el peso final (w_{ij}) de un término TF-IDF esta dada por:

$$w_{ij} = tf_{ij} \times idf_i \quad (6)$$

Consultas

Una consulta q es representada de la misma forma que la colección de los documentos. El término de los pesos w_{iq} de cada término t_i en q puede además ser computado de una forma similar a la de un documento normal. Como lo es:

$$w_{iq} = \left(0,5 + \frac{0,5f_{iq}}{\max \{f_{1q}, f_{2q}, \dots, f_{|V|q}\}} \times \log \frac{N}{df_i} \right) \quad (7)$$

Recuperación de documentos y Ranking de Relevancia

Muchas veces es difícil tomar la decisión de si un documento es relevante o no para cierta consulta. A diferencia del modelo Booleano, el modelo de vectores espaciales no toma esa decisión, ya que los documentos son rankeados de acuerdo a sus grados de relevancia con la consulta. Una forma de computar el grado de relevancia es calcular la *similitud* de la consulta q para cada documento d_j en la colección de documentos D . Hay muchas medidas de similitud. La más conocida y usada, la similitud coseno, la cual relaciona el coseno del ángulo entre el vector de la consulta q y el vector del documento d_j está dada por:

$$\cosine(d_j, q) = \frac{\langle d_j \bullet q \rangle}{\|d_j\| \times \|q\|} = \frac{\sum_{i=1}^{|V|} w_{ij} \times w_{iq}}{\sqrt{\sum_{i=1}^{|V|} w_{ij}^2} \times \sqrt{\sum_{i=1}^{|V|} w_{iq}^2}} \quad (8)$$

Esta similitud es también ampliamente usada en el clustering de texto y documentos. Otra función de similitud bastante conocida es el producto punto de dos vectores:

$$\sin(d_j, q) = \langle d_j \bullet q \rangle \quad (9)$$

Por otra parte, el ranking de los documentos se logra a través de sus valores de similitud. Los documentos mejor rankeados son catalogados como los más relevantes para la consulta.

También se puede obtener el grado de relevancia calculando directamente un puntaje de relevancia para cada documento de la consulta. El método **Okapi** y sus variaciones son ampliamente usadas para este propósito. La fórmula de recuperación de este método ([Hofmann, 2004],[Hofmann, 2001], [Huang et al., 2003],[Huang et al., 2004] y [Huang et al., 2007]) ha demostrado ser más efectiva que la fórmula del coseno para la recuperación de consultas cortas. Definiéndose el puntaje de relevancia Okapi de un documento d_j para una consulta q como:

$$okapi(d_j, q) = \sum_{t_i \in q, d_j} \ln \frac{N - df_i + 0,5}{df_i + 0,5} \times \frac{(k_1 + 1)f_{ij}}{k_1(1 - b + b \frac{dl_j}{avdl}) + f_{ij}} \times \frac{(k_2 + 1)f_{iq}}{k_2 + f_{iq}} \quad (10)$$

Donde se usa la siguiente nomenclatura:

- t_i : Un término de la consulta.
- f_{ij} : Contador de frecuencia bruta del término t_i en el documento d_j .
- f_{iq} : Contador de frecuencia del término t_i en la consulta q .
- N : Número total de documentos en la colección.
- df_i : Número de documentos que contiene el término t_i .
- dl_j : Longitud del documento d_j en bytes.
- $avdl$: Longitud de documento promedio en la colección.
- k_1 : Parámetro $k_1 \in [1, 2]$
- b : Parámetro b usualmente igual a 0,75
- k_2 : Parámetro $k_2 \in [1, 1000]$

Por último, otra función de puntaje conocida es la de la normalización de pesos pivoteada, denotada por pnw :

$$pnw(d_j, q) = \sum_{t_i \in q, d_j} \frac{1 + \ln(1 + \ln(f_{ij}))}{(1 - s) + s \frac{dl_j}{avdl}} \times f_{iq} \times \ln \frac{N + 1}{df_i} \quad (11)$$

Donde s es un parámetro usualmente igual a 0,2.

Hay que tener en cuenta que todas estas son funciones empíricas basadas en evaluaciones experimentales. En la práctica se suelen usar muchas variaciones de estas funciones.

2.4.3. Modelo de Lenguaje Estadístico

Este modelo, también llamado modelo de lenguajes, está basado en probabilidades y estadística. Para la recuperación de documentos primero se estima un modelo de lenguaje para cada documento y luego clasifica cada uno según la probabilidad de la consulta basada en ese modelo de lenguaje.

Sea la consulta q una secuencia de términos $q = q_1 q_2 \dots q_m$ y la colección de documentos D un conjunto de documentos $\{d_1, d_2, \dots, d_N\}$. Se considera además, que la probabilidad de que una consulta q sea generada por un modelo probabilístico basado en un documento d_j y

se denota por $Pr(q|d_j)$. Para clasificar a los documentos en la recuperación, interesa estimar la probabilidad inversa $Pr(d_j|q)$, que se obtiene usando la regla de Bayes:

$$Pr(d_j|q) = \frac{Pr(q|d_j)Pr(d_j)}{Pr(q)} \quad (12)$$

En estricto rigor, $Pr(q)$ no es necesario, ya que es la misma probabilidad para todos los documentos. Similarmente, $Pr(d_j)$ se considera uniforme, por lo que no afectaría la clasificación. Por lo tanto, lo único que es necesario computar es $Pr(q|d_j)$.

El modelo del lenguaje se basa usualmente en unigramas, es decir, solo se consideran términos individuales. Lo que se traduce en que el modelo asume que cada término es generado independientemente, lo cual representa una distribución multinomial sobre palabras, de la forma:

$$Pr(q = q_1q_2\dots q_m|d_j) = \prod_{i=1}^m Pr(q_i|d_j) = \prod_{i=1}^{|V|} Pr(t_i|d_j)^{f_{iq}} \quad (13)$$

Donde f_{iq} es el numero de veces que el término t_i aparece en q y $\sum_{i=1}^{|V|} Pr(t_i|d_j) = 1$. Lo que reduce el problema a estimar $Pr(t_i|d_j)$, que se puede calcular con la frecuencia relativa:

$$Pr(t_i|d_j) = \frac{f_{ij}}{|d_j|} \quad (14)$$

Recordando que f_{ij} es el numero de veces que el término t_i aparece en el documento d_j y $|d_j|$ el número total de palabras en d_j .

El problema de esta estimación es que cuando un término no aparece en d_j , se le asigna probabilidad 0, lo cual subestima claramente la posibilidad de ocurrencia de un término que no aparece en el documento. Este clásico problema se combate con una técnica llamada **suavizante**, la cual se basa en asignar probabilidades cercanas a cero, que muchas veces es más precisa que una probabilidad nula y de esta forma se logra tener una distribución más uniforme. Un suavizador aditivo clásico es de la forma:

$$Pr_{add}(t_i|d_j) = \frac{\lambda + f_{ij}}{\lambda |V| + |d_j|} \quad (15)$$

De forma que cuando $\lambda = 1$ se obtiene el **Suavizante de Laplace** y cuando $0 < \lambda < 1$ se obtiene el **Suavizante de Lidstone**.

CAPÍTULO 3

PROPUESTA DE SOLUCIÓN

En este capítulo se intenta explicar la solución propuesta para resolver el problema antes planteado. Se comienza dando detalles generales sobre la solución, la arquitectura del sistema y luego se entra en detalle en cada uno de sus componentes.

3.1. Diseño de la Solución

La solución como tal consiste de la re-implementación de *Inquiro.cl*², pero que este sea un motor de búsqueda vertical puramente chileno, es decir, que solo contenga dominios ".cl". Como se explicó anteriormente este buscador se diferencia de los que existen actualmente, como Google o Yahoo!, ya que estos se consideran buscadores horizontales, es decir, buscan sobre la web en general, sin ningún tipo de filtro por defecto.

Para lograr esto, se propone construir una colección representativa de la web chilena de por lo menos 10 millones de sitios web. Esto a su vez requiere la implementación de un *Crawler* o rastreador, el cual se encarga de extraer la información de la internet y recopilarla para su posterior uso en las búsquedas de términos.

La información a extraer de cada sitio se define como el texto que se encuentra en el cuerpo del documento, el título, la *url* y otros campos mas que se explicarán en detalle más adelante. Además, para reforzar la idea de búsqueda de contenidos, se incluyen los archivos de texto típicos que se suben a la web, como por ejemplo, los que presentan las extensiones .pdf, .doc, .xls y .ppt. Esto hace aun mas útil y diferenciador a *Inquiro.cl*, ya que los buscadores populares al ser generales, no son buenos buscando archivos en dominios acotados.

En la figura (4), se presenta la arquitectura inicial del sistema, el cual explica el flujo que recorre una consulta común en el buscador normalmente, desde que el cliente usa su navegador para acceder al sitio hasta que regresa a él con los resultados de la búsqueda y sus páginas correspondientes.

Se puede apreciar que el servidor *Inquiro* se compone principalmente de cuatro grandes componentes, un *crawler*, una aplicación *Django*, una aplicación *SolR* y su base de datos correspondientes.

La idea general del sistema es que el usuario ingresa términos para ser buscados, estos viajan a través de peticiones HTTP hasta el servidor principal de *Inquiro*. Una vez que la petición es reconocida como tal por la Aplicación *Django*, esta es procesada y se envía una consulta a la Aplicación *SolR*, la cual a través de su base de datos o índice invertido, aplica el algoritmo

²www.inquiro.cl

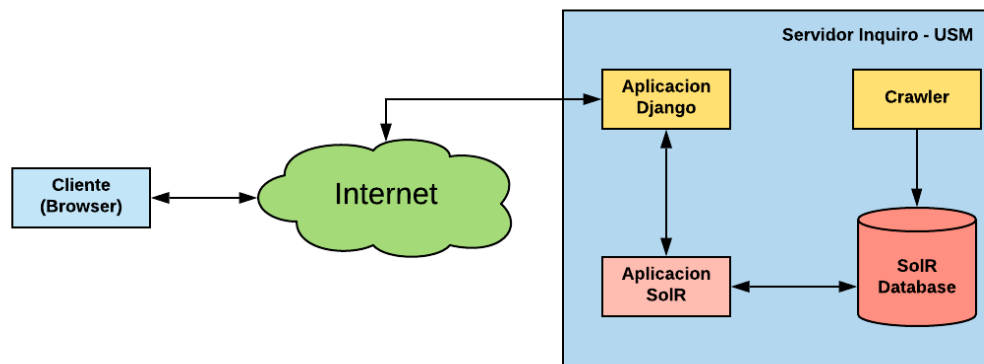


Figura 4: Diagrama Arquitectónico de Inquiro.cl
Fuente: Elaboración Propia.

de *Ranking* de **Vectores Espaciales**, específicamente aplicando el esquema de Frecuencia Inversa de Documentos (**TF-IDF**), explicado en el capítulo 2.

Una vez que el algoritmo es ejecutado y los documentos más relevantes son obtenidos, estos son enviados devuelta a la Aplicación Django, que es el encargado de mostrarle al usuario inicial los resultados de su búsqueda.

A continuación se procede a explicar la elección de tecnologías de cada componente.

3.2. Elección de Tecnologías

Este es el primer paso para la creación de un motor de búsqueda eficaz, escalable y valioso para aquellas personas interesadas en contenidos web chilenos. Es por esto que se realiza un detallado estudio de las mejores alternativas que ofrece la tecnología hasta el momento y que se presenta a continuación.

3.2.1. Crawler

Actualmente, según varios foros populares ^{3,4}, Scrapy⁵ y Heritrix⁶ son las herramientas de crawling preferidas hoy en día. Scrapy es un popular framework colaborativo y de código abierto que se especializa en la extracción de datos de páginas web [ScrapingHub, 2008]. Por otro lado, Heritrix es un proyecto de crawling web de código abierto, extensible y escalable

³<http://www.prowebscraper.com/blog/50-best-open-source-web-crawlers/>

⁴<https://www.quora.com/What-is-the-best-open-source-web-crawler-that-is-very-scalable-and-fast-And-why>

⁵<https://scrapy.org/>

⁶<https://webarchive.jira.com/wiki/spaces/Heritrix/overview>

[Heritrix, 2013]. Ambas herramientas tienen virtudes y defectos que se pueden comparar en la tabla 2.

| | Scrapy | Heritrix |
|---------------|---|---|
| Desempeño | Crawling eficiente. Escalabilidad depende de la implementación. | Crawling eficiente y escalable por defecto. |
| Espacio | Espacio en memoria no configurable | Espacio en memoria configurable. |
| Outputs | Almacenamiento en archivos CSV, XML y JSON | Almacenamiento comprimido de páginas en archivos Warc. |
| Documentación | Documentación clara y concisa. | Documentación escasa y compleja. |
| Configuración | Configuración total. Crawler depende de la implementación. | Difícil configuración de filtros y de las páginas semillas. |
| Parsing | Parseo se realiza directamente según implementación. | Parseo de los archivos Warc deben hacerse externamente |

Tabla 2: Comparación de características de crawlers populares.

Fuente: Elaboración Propia.

Como se puede observar, se estudian seis métricas relevantes para la elección del mejor crawler, las cuales son *Desempeño*, *Espacio*, *Outputs*, *Documentación*, *Configuración* y *Parsing*.

Al comparar el desempeño, es fácil evidenciar que Heritrix tiene la delantera, ya que por defecto es eficiente y escalable, es decir, no hacen falta muchas configuraciones para asegurar un buen desempeño del crawler. A diferencia de Scrapy, en el que el desempeño y escalabilidad dependen directamente de la implementación en código.

Para ver esto en la práctica, se realiza el siguiente experimento: Se ejecuta cada crawler por un tiempo fijo de un minuto de forma que rastreen la misma información desde un mismo sitio web y así poder comparar la cantidad de registros obtenidos y la cantidad de nuevos sitios alcanzados.

En este caso, el sitio web inicial escogido es "mercadolibre.cl", sitio web de compras, ventas y pagos por Internet muy popular en la actualidad. A continuación se presentan los resultados obtenidos en la tabla 3.

| | Scrapy | Heritrix |
|-------------|--|--|
| Nº Urls | 710 descargados y 122 encolados: 832 total | 153 descargados y 612 encolados: 765 total |
| Profundidad | 20 urls promedio | 10 urls promedio |
| Tiempo | 1m15s | 1m11s |

Tabla 3: Comparación de desempeño práctico de crawlers populares.

Fuente: Elaboración Propia.

Como se puede observar en esta tabla, ambos crawlers se comportan de manera similar, pero aun así existen diferencias importantes. Por ejemplo, en el número de urls rastreadas, *Scrapy* por poco posee mayor número de descargas que *Heritrix* en casi el mismo tiempo, pero lo interesante en este ámbito es que el primero descarga más sitios de los que encola y el segundo encola más sitios de los que descarga, lo que quiere decir, que *Scrapy* es más eficiente, ya que hace más trabajo del que acumula.

En cuanto a la profundidad promedio, *Scrapy* nuevamente tiene la ventaja con un valor de 20 urls en promedio, versus las 10 urls promedio de *Heritrix*. Esto también reafirma el paradigma del primer crawler, que es hacer más trabajo del que acumula.

Por estas razones, se puede concluir que *Scrapy* es más rápido en la práctica, ya que demora menos en procesar cada sitio rastreado.

En cuanto al espacio, *Heritrix* nuevamente tiene ventaja, ya que se puede configurar explícitamente el espacio en memoria a utilizar. *Scrapy*, por otro lado, al ser una librería en python, lenguaje de más alto nivel, no posee control de memoria explícito. La memoria se va llenando según las páginas encoladas en sus múltiples threads.

Los outputs o salidas, son una métrica importante en este caso, ya que más adelante se planea indexar la colección del crawler en el motor de búsqueda SolR y el tipo de archivos a importar debe ser compatible con este software. Ahora bien, si verificamos los tipos de archivos que *Heritrix* nos entrega, se puede apreciar que sólo entrega resultados del rastreo de páginas en archivos comprimidos llamados "WARC", que han demostrado un gran control y eficiencia en el manejo de la información de protocolos en aplicaciones en la red, como lo son HTTP, DNS, y FTP([ISO, 2009]). El problema es que SolR no acepta este tipo de archivos por defecto, lo que requeriría un pre-procesamiento adicional para indexar este tipo de colección. Por otro lado, *Scrapy* por defecto produce tipos de archivos CSV, XML y JSON, formatos mucho más populares que WARC, pero con mayor tamaño en disco.

Esto no es un problema, sino una ventaja, ya que SolR permite indexaciones justamente con estos tres tipos de archivos: CSV, XML y JSON. Ahora bien, si todos los tipos de archivo son soportados, claramente *Scrapy* tiene ventaja en este aspecto, pero falta decidir qué tipo de archivo de los tres sería el ideal en cuanto a espacio y eficiencia. Para verificar esto, se realizaron varias pruebas reales de crawling con *Scrapy* en un conjunto acotado de páginas, de forma que los outputs vayan variando entre CSV, XML y JSON cada vez. De esta forma, se comparan el tamaño de los archivos generados y el tiempo que demoraron en crearse, resultando en un claro ganador el formato CSV, puesto que registró el menor uso de espacio y el menor tiempo de creación.

La documentación también es un aspecto importante, ya que una buena documentación implica mejor entendimiento de la herramienta y por ende, mejores implementaciones para lo que se desea diseñar. Si comparamos ambas documentaciones, se puede apreciar que *Heritrix* no posee tanta documentación y tanta claridad como *Scrapy*, más bien es escasa y no es muy clara, lo que implica que no haya mucha gente que utilice y comparta sus implemen-

taciones, que a su vez produce que la información que se tenga sobre la herramienta sea mínima.

En cuanto a la configuración, ambas herramientas poseen diferentes formas de agregar opciones, por ejemplo, Heritrix basa la mayoría de ellas en el formato XML, es decir, para agregar o editar nuevas opciones hace falta configurar este tipo de archivos. Es más, no solo las opciones son en este formato, cada "Job" o "Trabajo de Rastreo", que es como se le llama a la forma en que se rastrea, también se configura de esta forma.

Esto se puede evidenciar al momento de tratar de aplicar filtros o al tratar de definir las páginas semillas para que comience a trabajar el crawler, funcionalidades básicas de todo rastreador, pero que en esta herramienta se vuelve bastante complejo. En este sentido, Scrapy posee la ventaja, ya que basa sus configuraciones y trabajos de crawling en el lenguaje de programación python, el cual aparte de que sea simple de diseñar, posee una gran flexibilidad, lo que es bastante importante a la hora de configurar lo que se desea extraer de la web con estas herramientas.

Finalmente, comparamos la manera en que estos softwares realizan la tarea de "parsing" o "parseo" de la información, es decir, como se filtra, discrimina o se extrae lo que se considera útil de cada texto sustraído de las páginas web seleccionadas.

Heritrix, por su parte, no posee un sistema de parsing propio, por lo que los archivos Warc que genera en su output deben ser parseados de forma externa, ya sea por un script o programa que permita hacerlo y que además pueda convertir este formato en uno compatible con los que posee SolR. En este sentido, Scrapy tiene ventaja, ya que es altamente flexible y convierte el proceso de parsing en una tarea muy simple y eficiente puesto que se realiza en el mismo ambiente del crawling, teniéndose la posibilidad de aplicar este parseo en tiempo real al mismo momento de la extracción de texto.

Claramente, por todas las comparaciones antes descritas, Scrapy se considera la mejor opción, ya que posee un gran sistema de parsing, sus formatos de outputs son compatibles con el sistema de indexación, su documentación es excelente y es altamente configurable.

3.2.2. Motor de Búsqueda

Este componente está muy ligado al anterior, ya que están estrechamente relacionados en el sentido de que el motor de búsqueda es el que tiene que indexar la información que extraiga el crawler escogido. En este caso, Lucene⁷ es de las API's de código abierto para recuperación de información mas populares actualmente, en esta API se basan herramientas aun más populares como lo son SolR⁸, motor de búsqueda de código abierto y Elasticsearch⁹, un servidor de búsqueda, que además provee un motor de búsqueda de texto completo.

⁷<http://lucene.apache.org/>

⁸<http://lucene.apache.org/solr/>

⁹<https://www.elastic.co/>

Teniendo estas herramientas en mente, se deciden comparar aquellas que posean características similares a SolR, por lo que se consideran los motores de búsqueda Lemur¹⁰ y Terrier IR¹¹, herramientas mas de ámbito experimental, pero que han demostrado tener un gran rendimiento.

Al tratar de comparar estas herramientas, lo primero que se realiza es revisar la documentación para verificar qué tan fácil y flexibles son para cumplir la tarea de indexar grandes cantidades de datos.

El primer software que se revisa es Lemur, su documentación no es muy clara y además está desactualizada, ya que hasta la fecha, su última actualización ha sido el 27 de Diciembre del 2016. Además, se verifica su configuración, y se puede apreciar que se basa principalmente en la sintaxis XML, tanto para la indexación como para la elección de parámetros, lo que limita un poco la flexibilidad de configuración.

La segunda opción que se verifica es Terrier IR, un motor de búsqueda escrito en Java y que es desarrollado por el departamento de informática de la Universidad de Glasgow en Reino Unido. Su documentación es más clara que Lemur y hasta la fecha, tiene una actualización reciente al 10 de Enero del presente año 2019.

Su configuración y diseño son bastante flexibles, ya que al basarse en el lenguaje de programación Java, permite un total control del flujo de la herramienta.

Finalmente, se chequea la documentación de SolR, motor de búsqueda escrito en Java con una arquitectura Cliente-Servidor, y se compara con las herramientas anteriores. Lo primero que se puede observar es que esta herramienta está siempre actualizándose, ya que hasta la fecha, su última actualización ha sido el 8 de Marzo del presente año 2019.

Su documentación está bien organizada y es sencilla de seguir, posee una configuración total de la herramienta, ya que provee una API que se adapta a los lenguajes de programación más populares como Python, Javascript y Ruby, entre otros.

Además, provee un sistema de administración con una interfaz gráfica simple de utilizar, que reemplaza a configuraciones XML, vistas en las herramientas anteriores.

Como la herramienta de crawling Scrapy ya ha sido escogida anteriormente, SolR tiene la ventaja por sobre las demás opciones, gracias a su gran documentación, a su API en Python y a su soporte de archivos con extensiones CSV, XML y JSON, que son justamente con las que trabaja Scrapy. Estas dos últimas razones, implican que el desempeño y simpleza del motor de búsqueda se vea favorecido.

¹⁰<http://www.lemurproject.org/>

¹¹<http://terrier.org/>

3.2.3. Framework Interfaz Gráfica

Este componente también cumple con una parte importante del buscador, que es la usabilidad de la interfaz gráfica, es decir, que tan fácil es poder buscar y encontrar las páginas web relevantes para cada consulta de los usuarios. Para esto, este componente debe poder conectarse de buena manera con los dos componentes antes escogidos, el crawler y el motor de búsqueda. Como Scrapy y SolR trabajan con el lenguaje Python, se investiga qué frameworks o ambientes de desarrollos web son compatibles con este lenguaje, por lo que se visualizan dos opciones populares: Flask¹² y Django¹³.

El primer framework, Flask, dice ser muy sencillo extremadamente rápido y su documentación es clara y práctica. La estructura de sus proyectos son libres, es decir, posee mucha flexibilidad de implementación, pero a la vez, esto puede significar en que todo lo que se necesite hay que instalarlo manualmente, ya sea desde lo más básico, como conexión con bases de datos o pruebas unitarias; hasta tareas más complejas como implementaciones de Api's REST.

Por otro lado, Django, posee una documentación clara y actualizada pero extensa, si bien esto puede ser bueno, a veces puede ser confusa con tanto para leer. La estructura de cada proyecto es generada automáticamente y provee herramientas de bases de datos, idiomas y testing unitarios integrados.

Tarda mas en renderizar que Flask, puesto que posee más funcionalidades por defecto, pero no deja de ser más robusto.

Por todas estas comparaciones, se selecciona a Django como la mejor opción, ya que facilita la implementación de todo lo que puede necesitar un sitio web, ya sea bases de datos, seguridad, configurabilidad y una gran cantidad de ejemplos en la red.

3.3. Diseño General de Componentes

Para el diseño de la solución se utilizaron dos lenguajes de programación orientados a objetos (OO): Python y C++, el primero es de más de alto nivel ya que su implementación se basa en el lenguaje C. El segundo es de más bajo nivel, ya que solo se diferencia con C, en que es orientado a objetos junto con variaciones en la sintaxis, pero ambos pueden controlar el espacio en memoria de forma eficiente.

Para la implementación del crawler se diseñaron dos scripts principales: crawler.py y Main.cpp. El primero, escrito en python, y como su nombre lo dice, se encarga de "crawlear" o "rastrear" páginas web, es decir, extraer todo el texto y los links que contenga cada páginas. El segundo, escrito en C++, se encarga de gestionar el comportamiento del crawler

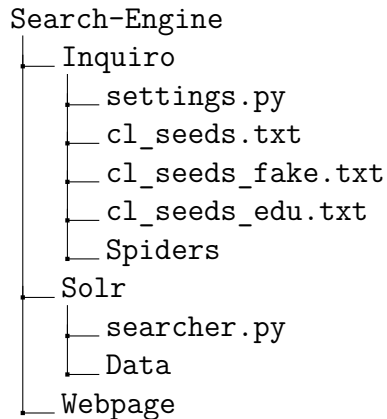
¹²<http://flask.pocoo.org/>

¹³<https://www.djangoproject.com/>

y automatizarlo con respecto a diferentes acciones y sus resultados.

3.3.1. Estructura General del sistema

La estructura general del sistema implementado posee todos los módulos, componentes y directorios necesarios para funcionar. Este fue diseñado para seguir la estructura propuesta en la figura (4) y que se presenta a continuación.



Como se puede observar, en un principio, el sistema cuenta con tres carpetas principales: *Inquiro*, *Solr* y *Webpage*. *Inquiro* contiene todo lo relacionado al rastreador o *crawler*. *Solr* está pensada para ser el contenedor del índice invertido del motor de búsqueda. Y *Webpage* contiene todo el despliegue del sitio web del buscador. A continuación se explica brevemente cada una de las funciones de los archivos y carpetas mencionadas.

- **Carpeta Inquiro:** Contiene todos los archivos necesarios para el buen funcionamiento del crawler diseñado.
- **Carpeta Solr:** Contiene todos los datos indexados por Solr para almacenar el índice invertido.
- **Carpeta Webpage:** Contiene todos los archivos necesarios para el despliegue del sitio web del buscador.
- **settings.py:** Archivo python que contiene las configuraciones generales de Scrapy. Lo más relevante de este archivo es la opción de ponerle nombre al rastreador, gestionar los *spiders* a usar y la opción de obedecer o no a los robots de cada sitio web.
- **cl_seeds.txt:** Archivo de texto que contiene la totalidad de Urls semilla. Estos sitios corresponden a un conjunto de más de 650 sitios recopilados de varios rankings y sitios populares de Chile en general.

- **cl_seeds_fake.txt:** Archivo de texto que contiene Urls de prueba, es decir, está pensado para usarse con fines de *testing*. Por defecto, el crawler usa este archivo como semillas iniciales.
- **cl_seeds_edu.txt:** Archivo de texto que contiene un subconjunto de *seeds.txt*, ya que está pensado para contener sólo las Urls de sitios más orientados a la educación y los contenidos de investigación.
- **Sub-Carpeta Spiders:** Contiene la implementación del *crawler* y todos los ficheros necesarios para su correcto funcionamiento. En el siguiente árbol de directorios se explica en detalle el contenido de esta carpeta.

El contenido de la sub-carpeta *Spiders* es de gran importancia por lo que requiere una explicación más detallada. A continuación se explica su contenido en el siguiente esquema:

```
Spiders
├── Outputs
├── Reoutputs
├── Spdlog
├── crawls
├── logs
├── crawler.py
├── recrawler.py
├── main.cpp
├── Makefile
├── visited.txt
├── revisited.txt
├── imported_files.txt
├── reimported_files.txt
└── graph.txt
```

Como se puede observar, las principales carpetas en *Spiders* son *Outputs*, *Reoutputs* y *Spdlog*. *Outputs* es sin duda el directorio con mayor importancia, ya que contiene toda la extracción de texto de todos los sitios web rastreados, en el formato de archivos CSV.

Reoutputs cumple una función bastante similar a la anterior, ya que contiene todos los archivos CSV del proceso de *Recrawling* o actualización de índices. Y *Spdlog* es la carpeta que contiene la librería C++ del mismo nombre, que permite utilizar mensajes de *logging*. A continuación se resume y explica todo el contenido de *Spiders* en detalle.

- **Carpeta Outputs:** Contiene todos los archivos CSV que se obtienen del proceso de *crawling* principal. Todos los archivos en esta carpeta poseen un nombre con el patrón *crawlN.csv*, siendo *N* el número de archivo actual, por ejemplo *crawl1.csv*, *crawl2.csv* y así sucesivamente.

- **Carpeta Reoutputs:** Muy similar a *Outputs*, contiene todos los archivos CSV que se obtienen del proceso de *recrawling* o actualización de índices.
- **Carpeta Spdlog:** Contiene la librería de *logging Spdlog*, que utiliza el *script main.cpp*.
- **Carpeta crawls:** Contiene el progreso de *crawling* de todo el sistema. Permite reanudar el proceso de rastreo de sitios, una vez que se ha detenido de manera segura.
- **Carpeta logs:** Contiene dos archivos de *logging*. *main_log.txt*, que informa todas las acciones que se ejecutan en el *script main.cpp* y *domains_log.txt*, que informa el dominio actual que se rastrea y cada cambio de dominio realizado.
- **crawler.py:** Archivo python que posee la implementación principal del *crawler*. Este archivo también es manipulado por el *script main.cpp*.
- **recrawler.py:** Archivo similar a *crawler.py*, contiene la implementación del *recrawler* o actualizador de índices. Se encarga de recorrer las *urls* del archivo *visited.txt* y re-visitarlas.
- **main.cpp:** Archivo C++ que posee la implementación general del sistema. Gestiona la ejecución de todos los comandos presentes en el Makefile y todos los archivos relacionados.
- **Makefile:** Archivo de sistema que contiene comandos simplificados para la ejecución de tareas complejas en el *script de main.cpp*. Un ejemplo de estos comandos son la ejecución del *crawler* o la importación de archivos CSV de la carpeta *outputs*.
- **visited.txt:** Archivo de texto que contiene todas las *urls* rastreadas actualmente por el *crawler*.
- **revisited.txt:** Archivo de texto muy similar a *visited.txt*, ya que contiene todas las *urls* que han sido re-visitadas y que fueron encontradas al actualizar las *urls* del archivo *visited.txt*.
- **imported_files.txt:** Archivo de texto que contiene el nombre de los archivos CSV de la carpeta *Outputs* que han sido importados al índice de Solr.
- **reimported_files.txt:** Archivo de texto muy similar al anterior, que contiene el nombre de los archivos CSV de la carpeta *Reoutputs* que han sido importados al índice de Solr.
- **graph.txt:** Archivo de texto que contiene una representación de grafo de los sitios rastreados. El formato de esta representación consiste en que cada línea de este archivo contiene dos url separados por un espacio en blanco. La url del lado izquierdo representa el origen del enlace y la url del lado derecho el destino del enlace entre estos dos sitios.

La última subcarpeta Webpage también es de gran importancia para el despliegue del sitio web del buscador, por lo que también requiere una explicación más detallada. A continuación se explica su contenido más importante en el siguiente esquema:

```
Webpage
├── Venv
├── Static
├── Search_engine
├── check_url.py
├── requirements.txt
├── uwsgi_params
├── webpage_nginx.conf
└── webpage_uwsgi.ini
```

Esta sub-carpeta, corresponde a un proyecto *Django*, es decir, contiene todos los módulos básicos de este framework web. Pero además de eso, presenta contenido adicional que se explica a continuación.

- **Carpeta venv:** Contiene la instalación local de las librerías python3 necesarias para el proyecto. Corresponde al entorno virtual creado localmente en cada máquina que ejecute el sitio.
- **Carpeta static:** Contiene todo el diseño del sitio web, como estilos, imágenes y colores.
- **Carpeta Search_engine:** Contiene la implementación principal del sitio. Los archivos *python urls.py* y *views.py* son los mas importantes. El primero es para la definición de rutas del sitio y el segundo corresponde a la implementación de la función principal del buscador, que es buscar en el índice invertido.
- **check_url.py:** Archivo *python* adicional. Se trata de un *script python* que permite visualizar resultados de la búsqueda en *Inquiro* en consola cuando el servidor web ya está en ejecución, y esto se logra a través de llamadas HTTP a la url del buscador. Por ejemplo, desde cualquier máquina se pueden ejecutar los siguientes comandos:

```
python check_url.py 10.6.32.122/search%3Dpdf%20usm/page%3D1
python check_url.py inquiro.cl/search%3Dpdf%20usm/page%3D1
```

Si todo sale bien, estos comandos deberían retornar el número de resultados y las url correspondientes al resultado de la páginas 1 de la búsqueda de documentos PDF de la USM, en consola.

- **requirements.txt:** Archivo de texto que contiene todas las librerías *python* necesarias para el correcto funcionamiento del sistema. Una vez que el entorno virtual está activo solo basta con instalar estas librerías con el siguiente comando linux:

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

- **uwsgi_params:** Archivo de texto plano que contiene los parámetros estándar para poder desplegar el *hosting* del sitio web del buscador con la ayuda de la librería *Uwsgi*.
- **webpage_nginx.conf:** Archivo de configuración del *hosting* de *Nginx*, que debe copiarse y enlazarse en las carpetas de instalación de *Nginx*.
- **webpage_uwsgi.ini:** Archivo de inicio del *hosting* de *Uwsgi*. Contiene la configuración necesaria para la creación de *Sockets* que permitan la conexión de máquinas externas al sitio web local.

3.3.2. Campos de extracción escogidos

Como bien se explicó en los capítulos anteriores, existen ciertas secciones de toda páginas web que son más importantes que otras, y este caso no es distinto. Es por esto que se extrae principalmente el texto de las secciones "título", "texto ancla", "cuerpo", "url" y "dominio" de cada sitio, además de los textos de ayuda: "tipo de documento", "contador" y "profundidad".

Todos estos campos le dan valor a la extracción de información de cada páginas web, por lo que un manejo adecuado de los datos es necesario en todo el proceso. El título y el texto ancla suelen ser textos bien acotados en longitud y por ende suelen tener una representación bien acertada del tema principal de cada documento. En cuanto al cuerpo del documento, este contiene la mayoría del texto de la páginas, es decir, provee del contenido mismo con las palabras o términos que pueden ser buscados posteriormente en las consultas.

La url representa el identificador único del sitio o sección de este, ya que es lo que nos permite acceder a una páginas web desde un navegador y es como la recuerdan las personas. El dominio en cambio, describe el nombre o el título de la máquina de destino que posee el sitio web y que se compone de varios tipos que se explican en detalle en las siguientes secciones.

También, se utilizan campos de ayuda adicionales a los anteriores, ya que poseen propiedades que ayudan a una mejor recuperación de documentos. El primero es el tipo de documento, que ayuda a evidenciar si un documento es del tipo HTML o un archivo con extensión. El segundo es un contador, que representa el orden en que fue rastreado ese sitio y que es meramente estadístico.

Y por último tenemos la profundidad del sitio, que representa que tan lejos se encontró en el proceso de *crawling* y que también es útil para estadísticas.

A continuación se explicarán en detalle las características principales de ambos scripts y como estos se interrelacionan para lograr un sistema automatizado y eficiente para la indexación del motor de búsqueda de esta memoria.

3.4. Diseño de Crawler.py

En primer lugar, hablaremos sobre el diseño del archivo python del crawler. Este posee una cantidad importante de componentes y librerías que se detallan a continuación.

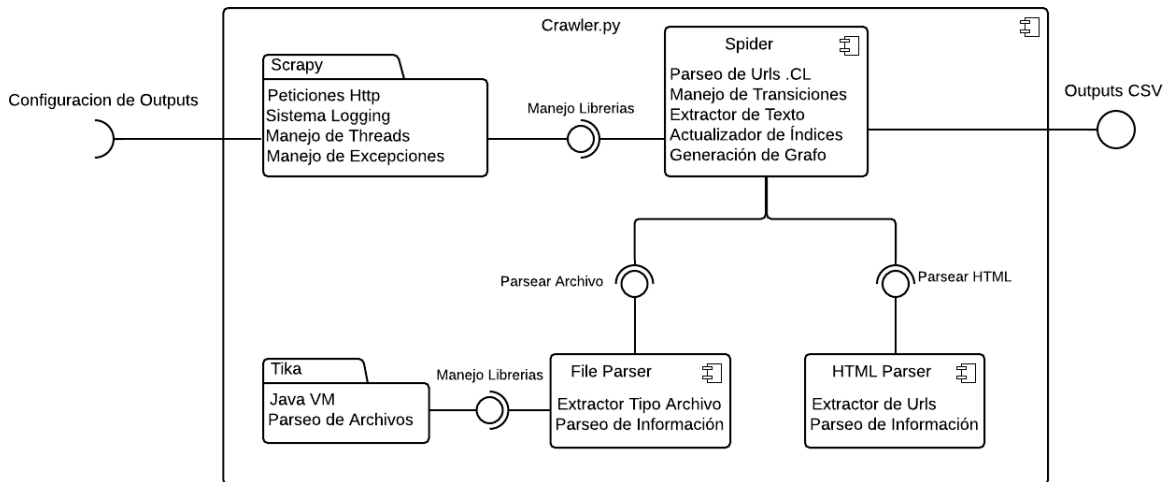


Figura 5: Diagrama de Componentes de script Crawler.py

Fuente: Elaboración Propia.

Como se puede observar en la figura (5), el componente Crawler.py, se compone de cinco componentes principales, un input de entrada y un output de salida. A continuación se procede a detallar cada elemento.

En primer lugar, tenemos el input de entrada, el cual se titula "Configuración de Outputs", esto se relaciona con la función principal del crawler, que es el de extraer todo el texto relevante de cada páginas web de Chile. Para esto, el crawler necesita que se le indique el nombre del archivo que contendrá el texto rastreado, el formato, ya que como se comentó en capítulos pasados, el formato puede ser XML, JSON y el que se va a ocupar, que sería el de CSV. Además, se pueden agregar opciones extras al rastreador, como por ejemplo, que se haga un rastreo silencioso sin mostrar texto del proceso que se realiza, o también se puede especificar si se desea crear un archivo tipo *checkpoint*, es decir, que pueda guardar el progreso del rastreo, ya sea guardando las variables utilizadas o guardando el manejo de colas de los sitios que esperan a ser rastreados.

3.4.1. Scrapy

Una vez que la configuración del output correspondiente es entregada al programa, esta información es procesada por la librería Scrapy, la cual se especializa en la extracción de información de páginas web y que es de mucha utilidad en este contexto, ya que gracias a este marco de trabajo, se pueden definir cuatro grandes funcionalidades principales:

- **Peticiones Http:** A diferencia de otras librerías, Scrapy permite realizar peticiones o accesos a páginas web de forma eficiente, de forma que en una sola *request* o petición, se tenga a disposición toda la información disponible de la páginas visitada.
- **Sistema Logging:** Si bien esta librería se especializa en el rastreo de información, también está pensada para poder trabajar en servidores remotos, por lo que necesariamente debe tener un sistema de *logging*, es decir, un sistema que permita visualizar qué es lo que hace el sistema en todo momento y que quede registrado en archivos llamados "logs", que contienen una gran variedad de información con respecto a cada tarea realizada y que además son altamente configurables para todo tipo de necesidades.

$[dd/MM/YYYY - HH : mm : ss][Tipo Log] Descripción$ (16)

En este caso, como se puede observar en (16), lo más relevante del logging implementado es la fecha y hora, tipo de logging, y una pequeña descripción de lo que se hace para cada tipo de acción. Para el tipo de logging se definen 4 tipos en inglés: info, debug, warning y error. El primero solo tiene la intención de informar sobre la acción. El segundo se utiliza más para cuando se necesitan verificar variables o corregir errores en cierta parte del sistema. El tercero se utiliza para cuando se requiere alertar de algo, por ejemplo, en este crawler en particular se utiliza para cuando el sistema está esperando algo o cuando se inicia y apaga el sistema.

Por último, el tipo error, como su nombre lo dice, se utiliza cuando se producen errores generales en el sistema, de forma de poder registrarlos e identificarlos rápidamente y a la vez poder encontrar la causa.

Ahora bien, todo esto es utilizado con un fin bastante simple, el cual es hacer un seguimiento del rastreo de dominios actuales. En otras palabras, se utiliza el sistema de logging principalmente para saber cual es el dominio actual al cual se está extrayendo información y saber en qué fecha y hora ocurren cambios de dominio.

Todos estos mensajes *logging* son almacenados en el archivo `domains_log.txt`, específicamente en el directorio *Logs* que se encuentra en los *spiders* del proyecto.

- **Manejo de Threads:** Scrapy se caracteriza por ser de los rastreadores más rápidos y eficientes, y esto lo logra a través de un manejo de "threads" o hilos excepcional, que permiten procesar múltiples sitios web al mismo tiempo, de forma que la información obtenida sea coordinada y sincronizada con la cola de sitios que se van encontrando en el proceso. En otras palabras, cómo cada sitio web posee más links a otros sitios, estos se van guardando en una cola y luego los threads se encargan de rastrear cada páginas guardada coordinadamente.
- **Manejo de Excepciones:** Todo programa sin excepciones puede tener errores, y es importante saber manejarlos de forma responsable. En este sentido, Scrapy no se queda atrás, ya que posee un sistema de manejo de excepciones bastante avanzado, de forma que permite definir y configurar los tipos de errores que se quieren manejar según sean las necesidades.

En el caso de este crawler, se utilizan para cuando se encuentra un sitio que responde

con un código de error 404, el cual significa que no se encuentra el sitio solicitado. Por otro lado, también se utiliza para cuando se ha sobrepasado la cantidad máxima de sitios rastreados por archivo, provocando que se cierre la escritura en el archivo actual y que empiece una nueva escritura en un archivo diferente.

Como se puede observar, Scrapy posee un gran entorno de trabajo para trabajar con la información en la web. Pero esto no es todo, estas funcionalidades que se explicaron son solo las que se consideraron esenciales para el proyecto y que son bien manejadas por el siguiente componente que se detalla a continuación.

3.4.2. Spider

El componente Spider, es el próximo a detallar, y es en pocas palabras, la clase que es definida por Scrapy para gestionar todas las necesidades del usuario en cuanto a la extracción de información de páginas web. Entre estas necesidades, se destacan cuatro principales, que son las más relevantes para el proyecto, y que se explican a continuación.

- **Parseo de Url's chilenas:** Este es uno de los requerimientos esenciales de la memoria. Es el que le da sentido a todo el proyecto, ya que se trata de la comprobación de si un sitio pertenece a la web chilena o no. Para lograr esta comprobación, existen varias técnicas, unas más eficientes que otras evidentemente. Por lo que se opta por una opción simple y eficaz, las expresiones regulares. Estas expresiones regulares según [Goyvaerts, 2017] se definen como secuencias de caracteres que forman un patrón de búsqueda, principalmente utilizadas para la búsqueda de patrones de cadenas de caracteres u operaciones de sustituciones. Estas pueden llegar a ser muy útiles si se utilizan de la manera correcta. En este caso, para lograr el objetivo de identificar las url chilenas, se diseña una expresión regular capaz de identificar, en primer lugar, como es la forma de una url general y en segundo lugar, que pertenezca al dominio ".cl", es decir, que tenga la siguiente forma:

http : //www.dominio.cl/ruta/index.html (17)

Como se puede observar en la ecuación (17), la url a identificar puede contener hasta cuatro secciones principales ([educativa Partesdel.com, 2018]), las cuales son las siguientes.

- **Protocolo:** Es la parte de la URL que se conoce como el formato en el cual se intercambian las informaciones de la máquina de destino y del navegador que se está utilizando. Sección azul en (17). Existen distintos protocolos que se clasifican tomando en cuenta la clase de contenido que se quiere intercambiar. **Http** es el protocolo principal y se utiliza para

visualizar páginas web. **Ftp**, en cambio, se utiliza para transferir archivos. Por último, **Mail** es utilizado para el traslado de correos electrónicos. Al igual que los dos puntos, las barras inclinadas ("/") sirven de separación y son las que inician el contacto con la dirección de Internet para lograr la búsqueda del servidor.

- **Servicio:** Esta sección, de color rojo en (17), se suele omitir algunas veces y es muy común que se utilice el servicio **www**, conocida como la "red informática mundial", ya que es el sistema lógico de acceso y búsqueda que utilizan los sitios web en general.

- **Dominio, Tipo de Dominio y País:** La siguiente sección, la de color verde en (17), se puede llamar también Servidor, ya que describe el nombre o el título de la máquina de destino que posee el sitio web. El nombre es muy importante, ya que es la representación directa del sitio al cual se está accediendo. Después del nombre, viene el dominio de nivel superior, que corresponde a las tres últimas letras que tiene el nombre del servidor. Estas permiten identificar el tipo de sitio, además del país de procedencia.

Existen dos tipos de dominios: Geográficos o genéricos. Siendo el más común y más utilizado el genérico ".com". Los genéricos se aplican según la actividad que realicen. Por ejemplo, ".com" es comercial, ".edu" se refiere a educación, ".org" representa entidades gubernamentales o sin fines de lucro, entre otros. Por otro lado, los geográficos o territoriales, como su nombre lo indica se refiere al territorio donde se encuentra la dirección web. Algunos ejemplos son: .ec (Ecuador), .ar (Argentina), .es (España), .cl (Chile), entre otros. Aunque esto no es una obligación, ya que en ocasiones, servidores de ciertos países se registran con dominios genéricos sin ningún problema.

Además, se puede dar el caso en que el dominio contenga subdominios, estos se utilizan cuando se requiere atender con diferentes nombres en una misma máquina. Se suelen colocar en la parte delantera del nombre de la máquina. Un ejemplo de esto es que el nombre del servidor tenga como dominio "musicaenlaweb", el nombre de la máquina puede ser "blogspot" y el dominio de nivel superior ".com".

Por último, también se puede especificar el puerto de la máquina, el cual es un número correspondiente a un punto determinado de entrada a la máquina. El protocolo "http" normalmente utiliza el puerto 80. Por esta razón es que no se especifica al colocarse la URL. En algunas ocasiones, en algunos navegadores cuando se coloca la URL, el puerto se oculta. En el caso del FTP, el puerto es el número 21.

- **Ruta y Nombre de archivo:** La ruta, color morado en (17), hace referencia al orden de subsecciones que tenga el sitio o también al listado de carpetas que la computadora necesita acceder para encontrar los archivos del sitio. En este sentido, las rutas de los sitios son similares a los directorios que tienen los sistemas operativos. El directorio se utiliza para mantener el orden de las páginas que se tienen en el servidor y su estructura se caracteriza por ser jerárquica. Además de que existe la posibilidad de anidar unos directorios en otros.

- **Manejo de Transiciones:** Esta es otra de las necesidades fundamentales de este componente, ya que para poder cumplir la tarea de rastrear toda la web chilena, se debe hacer con toda la eficiencia posible, ya sea paralelizando el trabajo con la ayuda de threads simultáneos o con el paso de un sitio a otro. Es por esto que se requiere de un buen sistema de transiciones, es decir, saber coordinar de manera eficaz el paso simultáneo de un sitio a otro, ya que existe información valiosa que es transferida entre estos, como lo es el caso del **texto ancla**.

Este término hace referencia al texto que aparece en los "links" de los sitios que pueden apuntar a otro sitio o a otra sección del mismo. Es por esto que se necesita que una vez que el rastreador encuentre muchos links a otros sitios, se guarde el texto ancla correspondiente y se transfiera esa información a la siguiente recursión del Spider que rastree texto perteneciente al sitio encontrado. En otras palabras, se transfiere el texto ancla de una instancia del rastreador a otra, de forma que este texto se guarde junto a la demás información encontrada, como lo es el texto del cuerpo del documento, el título, entre otros.

- **Extracción de texto:** Esta funcionalidad también es de las más necesarias ya que en esta recae la responsabilidad fundamental de extraer la información como tal de cada sitio. El Spider está diseñado para esta labor, y es por esto que junto a Scrapy se tiene a disposición los más eficientes algoritmos de extracción.

La extracción de texto se relaciona fuertemente con el concepto del parseo, de hecho se define una función "parse()" que se encarga de parsear y extraer el texto de cada sitio a través del manejo del *response* o respuesta que entrega cada páginas al solicitar acceso a ella a con una *request* o petición que en este caso es gestionada por Scrapy internamente.

- **Actualizador de Índices o Re-Crawling:** El actualizador de índices, que consta principalmente de la acción de re-rastrear sitios web que ya han sido rastreados, con la finalidad de obtener la versión más reciente de cada documento. Esto está pensado para realizarse cuando algunos sitios rastreados alcancen un cierto tiempo de indexación.

Si bien, esta acción se puede considerar como otra instancia *Spider*, la finalidad es bastante similar a la extracción normal de texto de páginas web, pero hay una diferencia importante entre estos, y es que este último solo rastrea los sitios ya visitados. En otras palabras, solo se encarga de recorrer los *links* del archivo *visited.txt*, y extraer el texto de ellos, pero no almacena las urls de cada sitio revisitado para seguir con el crawling.

- **Generación de Grafo:** Finalmente, se tiene la funcionalidad de generar un grafo, ya que a partir del rastreo que se produce de un sitio web a otro, es posible almacenar el *path* o camino utilizado para llegar a cada páginas. Siendo cada una de estas un nodo y la unión de estos un enlace o *link* entre estos sitios.

La representación implementada se trata de un archivo de texto plano llamado *graph.txt*, muy similar a *visited.txt*, pero que almacena cada url de una páginas visitada, y junto a ella la url de una páginas a la que apunta, es decir, cada línea de este

archivo contiene dos url separadas por un espacio, que representan la unión. Este grado de información de enlaces puede servir para futuras investigaciones de la red chilena y poder dimensionar de mejor forma como se comporta Chile en la internet.

Con estas funcionalidades, la desafiante tarea de obtener una colección representativa de la web chilena se hace mucho más simple. Pero eso no es todo, en relación con la funcionalidad de extracción de texto, se entrará en más detalle a continuación en qué es lo que se extrae exactamente y en los tipos de extracción que se realizan.

Para empezar, hay que tener muy claro cuáles serán los tipos de extracciones que se harán, o mejor dicho, desde que formato se va a extraer la información.

Por temas de diseño, este proyecto se centra únicamente en la extracción de texto, es decir, solo en formatos que soporten texto. Entre los cuales se pueden dividir en dos grandes componentes, los documentos HTML como tal y los Archivos.

3.4.3. HTML Parser

El Parseador de HTML, como se mencionó anteriormente, corresponde a un subcomponente del componente de Extracción de Texto, y se diferencia de los demás subcomponentes debido a su comportamiento particular con los documentos web. En primer lugar, este parseador realiza dos tareas fundamentales y únicas en su tipo que se explican a continuación.

- **Parseo de Información:** Como bien se ha explicado a lo largo de esta memoria, los documentos HTML poseen información valiosa en forma de texto, y que si bien todos estos documentos son diferentes y poseen una variedad infinita de elementos, estos aun así comparten una estructura común, ya que HTML es un lenguaje de marcas estándar y transversal para todo sitio web. Ver figura (6).

Esta estructura en común consta de cinco campos de texto principales para identificar un sitio web. El cuerpo, el título, la url, el dominio y el texto ancla. Estos campos fueron convenientemente escogidos por la relevancia que poseen para una búsqueda efectiva de términos y que evidenciamos a continuación.

Body

El *body* o "cuerpo" de un documento web corresponde a la sección más importante, ya que como se puede observar en la figura (6), es donde se encuentra el contenido de la páginas como tal. Esta puede contener un sin número de elementos, como lo son títulos, párrafos, imágenes, documentos, entre otros, y que dependiendo del sitio son estructurados de diferentes formas, es decir, presentan la información a gusto del diseñador.

Dada la importancia de esta sección, es fácil imaginar que esta se trate de una de grandes proporciones en general, y por ende, la que nos provee la mayor cantidad de

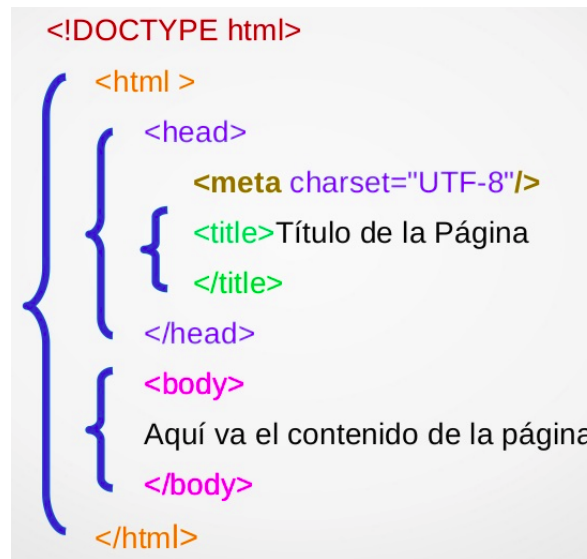


Figura 6: Estructura estándar de documentos HTML

Fuente: [Navia, 2015].

palabras o términos en forma de texto para el buscador.

Pero esto no es todo, en el cuerpo además podemos encontrar links o urls que apuntan a otros sitios web, lo cual es lo más importante para un crawler. Hablaremos más de esto en la funcionalidad "Extracción de Urls" de este mismo subcomponente.

Título

Como su nombre lo dice, se trata del título del documento, es decir, el texto que típicamente aparece en las pestañas de los navegadores. Es definido en la cabecera (*head*) de cada documento HTML y es de gran importancia para el proyecto, ya que define extremadamente bien el contenido del sitio. Es por esto que al momento de realizar búsquedas de términos, se le da más importancia a las palabras que aparecen en los títulos que a los que aparecen en el cuerpo.

Url

Más que una sección, la url propia de un sitio es lo que le da la identidad como tal. Es lo que las personas en general más recuerdan de un sitio y es que además es el punto de acceso a ella, es decir, teniendo este link se puede visualizar el sitio web completo en cualquier navegador.

Dominio

Si bien, el dominio de cada sitio web puede considerarse como una subsección de la Url, este también es almacenado independientemente, de forma que se refuerce el ranking de los documentos que posean los términos buscados en el dominio.

Esto también ayuda a la hora de hacer estadísticas, ya que se pueden añadir filtros simples para contar cuantos subdominios tiene cierto dominio.

Texto ancla

Este campo de texto es mucho menos conocido que los demás antes mencionados, pero igual de importante. El texto ancla corresponde al texto presente en los links del cuerpo de las páginas visitadas. Muchas veces contienen textos indicativos como "Haga click aquí" o "Comprar", que si bien tienden a ser textos cortos, muchas veces son de gran ayuda para describir el sitio al que enlazan.

El problema de este campo, es que no se encuentra en el sitio web actual, ya que corresponde al texto qué otro sitio tiene para describir el actual, por lo tanto este debe transmitirse desde el sitio de origen al actual para ser parte de los términos indexados del sitio web en cuestión. De esto se encarga el crawler, de transmitir fielmente cada texto ancla recursivamente a cada sitio nuevo visitado.

Campos adicionales (Tipo de Documento y Fecha)

También se incluyen campos adicionales, que no derivan directamente de documentos HTML como tal, pero son de gran ayuda a la hora de buscar términos de consultas, ya que refuerzan el ranking obtenido de cada documento que contenga relación con estos campos. Estos son el Tipo de Documento rastreado y la Fecha de rastreo.

El primero tiene que ver con la denominación que se le da a cada sitio visitado, ya que como vimos en la figura (5) se hace diferencia entre documentos HTML como tal y documentos de archivos como PDF, DOC, XLS, entre otros. Gracias a esto, la tarea de buscar documentos con cierta extensión se vuelve simple, ya que basta con poner la extensión de los archivos a buscar como término de búsqueda. Por ejemplo, si se necesita buscar solo archivos educativos con la extensión PDF, se puede lograr fácilmente con la consulta "educación pdf".

Por otro lado, la fecha de rastreo también cumple un rol importante, ya que es el campo que indica si un documento es reciente o no, y por lo tanto, el que indica si hace falta una actualización. Este campo también es vital al momento de realizar el *ranking* de documentos según una consulta de términos, ya que además de ordenar por el puntaje de relevancia que posee cada documento, se ordena también por fecha, es decir, cuando dos documentos contengan puntajes similares, el más reciente tomará más relevancia y por ende, será más visible que el otro.

■ Extracción de Urls:

Además de extraer información, el parseador HTML tiene un rol fundamental que no posee ningún otro subcomponente, y es el de extraer urls o links de las páginas web que se visitan, es decir, al rastrear el cuerpo de cada documento HTML se pueden encontrar links o urls que apuntan a otros sitios web, lo cual es lo más importante para un crawler o rastreador, ya que estas urls se van encolando de forma de poder seguir con el rastreo en esas páginas. En otras palabras, el crawler necesita encontrar nuevos links en el *body* de cada sitio, de forma de ir avanzando en cadena hacia ellos para extraer información.

3.4.4. File Parser

A diferencia del subcomponente anterior el File Parser o "parseador de archivos" posee roles bastante diferentes, y es que se trata del parseo de documentos que estan en línea, pero que no corresponden a un sitio HTML como tal, es decir, son todos los documentos de uso común para las personas en sus diferentes formatos. A continuación se detalla las dos principales funcionalidades de este subcomponente.

■ Extracción de los Tipos de Archivo

En primer lugar, para poder parsear correctamente cada archivo, hace falta conocer el formato en que este se encuentra. Ya que cada uno posee una forma distinta de representar la información.

Para lograr esto, se hace uso de las expresiones regulares, ya que es fácil evidenciar que cada url de este tipo de documentos termina justamente con la extensión que corresponde a cada formato. Por ejemplo, en el caso de los archivos PDF, su extensión es ".pdf", en el de los archivos Microsoft Word, su extensión es ".doc", en el de los archivos Microsoft Excel, su extensión es ".xls" y en el de los archivos Microsoft PowerPoint, su extensión es ".ppt" y así sucesivamente.

En el caso de esta memoria, se rastrean los tipos de archivo más utilizados en la web y que justamente son los mencionados anteriormente. En pocas palabras, las extensiones que se utilizan para indexar en el buscador son las siguientes: '.pdf', '.doc', '.docx', '.xls', '.xlsx', '.ppt' y '.pptx'.

Como se puede observar, varios de estos archivos son similares entre sí, pero algunos presentan una *x* al final de su extensión. Esto ocurre solo en los archivos Microsoft¹⁴, ya que las extensiones mas antiguas y estándar usaban estas versiones ("docx", "xlsx" y "pptx") en los años 2007.

A pesar de su antigüedad, estos se encuentran en gran cantidad en la web, por lo que es conveniente incluirlos en el rastreo de la información para el buscador.

También es importante saber que en este subcomponente se realiza la captura de extensiones ".zip", ".jpg" y ".png", pero como estos tipos de archivo no corresponden a texto como tal, son descartados. De esta forma, se evita que cualquier tipo de extensión nueva o desconocida sea tratada como HTML, y por ende se evitan dobles comprobaciones a la hora de parsear.

■ Parseo de Información

A diferencia del parseo de información de los documentos html, este subcomponente tiene una forma muy simple y a la vez muy particular de extraer la información, ya que una vez que el tipo de archivo es detectado, se hace uso del siguiente componente externo *Tika*¹⁵ para la extracción eficiente de texto. A continuación, se explica esto en detalle.

¹⁴<https://www.microsoft.com>

¹⁵<https://tika.apache.org/>

3.4.5. Tika

Este es un componente externo pero muy relevante para el componente File Parser antes mencionado, ya que corresponde a una librería Apache¹⁶ externa escrita en Java, pero que posee una API escrita en python, que es la que se utiliza para este proyecto.

Esta librería es experta en detectar y extraer metadatos y texto de más de mil tipos de archivos diferentes. Todos estos tipos de archivos se pueden analizar a través de una única interfaz, lo que hace que Tika sea útil para la indexación de motores de búsqueda, análisis de contenido, traducción y mucho más.

A continuación se presentan las dos principales funcionalidades que Tika ofrece para ayudar a construir un buen motor de búsqueda.

- **Java Virtual Machine (JVM)**

Segun ([Venners, 1996]) La *Java Virtual Machine* (JVM) o "máquina virtual de Java", es una computadora abstracta que ejecuta programas compilados de Java. La JVM se dice que es "virtual" porque generalmente se implementa en software sobre una plataforma de hardware "real" y su sistema operativo. Todos los programas de Java se compilan para la JVM.

Como Tika está implementado en Java, es natural que su ejecución se realice en la JVM, pero lo particular de esta funcionalidad es que le permite a Tika funcionar como una estructura Cliente-Servidor, es decir, Tika provee un servidor local, al cual el usuario puede realizar consultas, siendo esas consultas, en este caso, la extracción de texto de diferentes tipos de archivos con sus formatos particulares para poder procesarla fácilmente.

- **Parseo de Archivos**

Aquí es donde se realiza el parseo real de los archivos que se encuentran en el sub-componente anterior. Tika posee algoritmos específicos para la extracción eficiente de cada tipo de archivo soportado. Es por esto que en la implementación de este parseador, solo hace falta definir el objeto parseador Tika y luego indicarle la ruta del archivo en cuestión, ya sea una ruta web o una ruta local de un computador. De esta forma, se obtiene solamente el texto relevante de cada uno de estos archivos.

Finalmente, luego que el componente Spider extrae y procesa la información de cada sitio web, este retorna los resultados en forma de archivos de texto CSV, los cuales justamente corresponden a los outputs de este gran componente mayor, que en realidad corresponde al script llamado Crawler.py, es decir, el rastreador general.

Pero eso no es todo, luego estos outputs son utilizados y gestionados por el siguiente gran componente del que hablaremos a continuación.

¹⁶<https://www.apache.org/>

3.5. Diseño de Main.cpp

Ahora que ya hemos hablado del Crawler, uno de los componentes más importantes en el sistema, ahora toca hablar como este es usado y gestionado por este nuevo componente llamado Main.cpp, que planea estar a un nivel más alto de abstracción, es decir, tiene un control más grande de los recursos del sistema.

A continuación, hablaremos sobre los componentes y librerías presentes en Main.cpp y que se resumen en la figura (7).

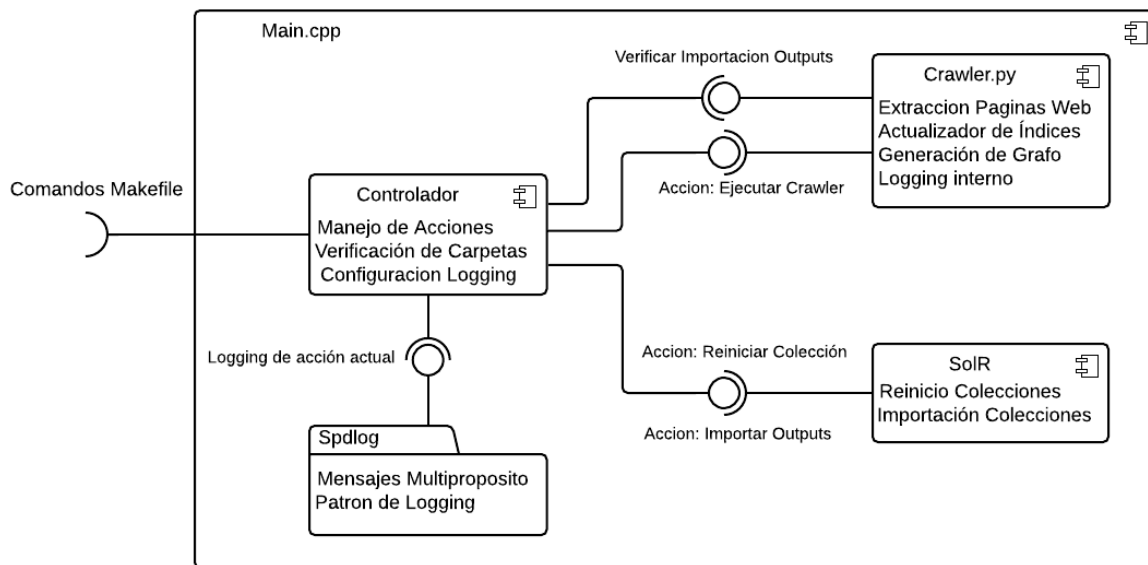


Figura 7: Diagrama de Componentes de script Main.cpp

Fuente: Elaboración Propia.

Como se pueden observar, el componente Main.cpp contiene a su vez cuatro componentes principales y un input o entrada. A continuación se procede a detallar cada elemento.

En primer lugar, llama la atención de que este componente solo tenga una entrada y no una salida. Esto se debe a que Main.cpp está diseñado de forma que gestione automáticamente todos sus subcomponentes, y que algunos de estos produzcan salidas. Las que luego también son auto gestionadas y así sucesivamente.

Para explicar la única entrada de este componente, primero se debe explicar que es un *Makefile*. Según ([Garver, 1998]) Make es la utilidad del sistema operativo Unix que está diseñada para iniciar la ejecución de un *Makefile*. Este se trata de un archivo especial, que contiene comandos de *shell* a ser ejecutados sucesivamente una vez que se ejecuta el comando Make.

Justamente el único input que existe se refiere a esto, los comandos del archivo *Makefile* creado, y que en general se resumen en siete comandos en inglés: *help*, *build*, *run*, *restart*

solr, *import files*, *restart and import* y *restart and run*. Comandos que se detallan a continuación en el manejo de acciones del Controlador.

3.5.1. Controlador

El controlador es el primer componente que tiene contacto directo con los comandos *Makefile* ingresados por el usuario y es el responsable de redirigir y ejecutar las diferentes acciones que ofrece el programa por separado. Se procede a explicar las tres funcionalidades principales del controlador.

■ Manejo de Acciones

Típicamente, esta es la funcionalidad principal de todo controlador, pero en este caso en particular es un poco diferente. El manejo de acciones se hace a través de la ejecución de comandos complejos a partir de comandos *Makefile* más simples. En la tabla (4) se resumen todos los comandos disponibles y que permiten manipular el sistema según lo que se necesite.

| Makefile | Terminal |
|---------------------------|---|
| <i>help</i> | @perl -e '\$(HELP_FUN)' \$(MAKEFILE_LIST) |
| <i>build</i> | g++ main.cpp -o main —std=c++17 —lstdc++fs -I"spdlog/include" |
| <i>run</i> | ./main —run |
| <i>recrawl</i> | ./main —rec |
| <i>restart_solr</i> | ./main —res |
| <i>import_files</i> | ./main —imp |
| <i>reimport_files</i> | ./main —rimp |
| <i>restart_and_import</i> | ./main —res_and_imp |
| <i>restart_and_run</i> | ./main —res_and_run |
| <i>full_clean</i> | rm -rf outputs/*.csv rm -rf crawls/* rm -rf logs rm -rf imported_files.txt rm -rf visited.txt rm -rf *.pyc |

Tabla 4: Comandos Makefile para ejecutar en el Controlador.

Fuente: Elaboración Propia.

Al analizar esta tabla, es fácil visualizar que los comandos Makefile de la primera columna son mucho más entendibles que los de la derecha. Esto se debe a que estos últimos contienen comandos de compilación para programas C++, además de

opciones que cambian el comportamiento del programa en sí. A continuación se explica cada comando en detalle.

Ver lista de comandos disponibles

`$ make help` (18)

Este comando sirve para ver todos los comandos disponibles en el MakeFile. Despliega un mensaje explicando la función de cada uno. Ver figura (8).

```
francisco@ubuntu-18: ~/Escritorio/Github/Search-Engine/inquiro/spiders$ make help
usage: make [target]

miscellaneous:
  help      Para mostrar esta ayuda

options:
  build      Compilar el script general (necesario para los demas comandos)
  run        Correr el crawler sin reiniciar la coleccion
  recrawl    Correr el recrawler que revisitará las url de visited.txt
  restart_solr  Reiniciar la coleccion SolR
  import_files  Importar archivos csv de la carpeta outputs
  reimport_files  Re-Importar archivos csv de la carpeta reoutputs
  restart_and_import  Reiniciar SolR e Importar (2 comandos anteriores en secuencia)
  restart_and_run  Reiniciar SolR y Correr el crawler (sin importar outputs)
  full_clean  ATENCION! BORRARA TODOS LOS OUTPUTS Y REGISTROS! (HARD RESET)
```

Figura 8: Mensaje desplegado por el comando help del Makefile

Fuente: Elaboración Propia.

A continuación se explican los comandos reales del MakeFile.

Compilar script principal main.cpp

`$ make build` (19)

Este comando ejecutará el siguiente comando de compilación necesario para el archivo main.cpp:

```
g++ main.cpp -o main --std=c++17 -lstdc++fs -I"spdlog/include"
```

Se trata simplemente de una compilación C++ del archivo main.cpp indicando un nombre de ejecutable main, pero que además requiere de ciertas opciones para funcionar, como lo es la compilación de la versión 17 de la librería std, que hace referencia al manejo avanzado de los strings.

También, se utiliza la opción lstdc++fs, que corresponde a la compilación de librerías que permiten acceder a los directorios locales y manejarlos eficientemente.

La última opción hace referencia al input de la librería externa spdlog, la cual brinda un amplio poder de logging en C++.

Ejecución Crawler Principal

`$ make run` (20)

Este comando ejecuta el crawler principal, es decir, ejecuta el archivo Main.cpp en modo crawler. Ver figura (9).

```
francisco@ubuntu-18:~/Escritorio/Github/Search-Engine/inquiros/spiders$ make run
./main --run
RASTREANDO NUEVAS PAGINAS
No handlers could be found for logger "scrapy.utils.log"
-----
('Con: ', 0, 'Depth: ', 0, 'http://www.colegiocastelgandolfo.cl', '<', '', '')
-----
('Con: ', 1, 'Depth: ', 1, 'https://www.myschool.cl/indexMS.php?co=Castelgandolfo', '<', 'http://www.colegiocastelgandolfo.cl', u'Ingre
sistema')
-----
('Con: ', 2, 'Depth: ', 1, 'http://www.colegiocastelgandolfo.cl/admision/', '<', 'http://www.colegiocastelgandolfo.cl', u'Admisi\xf3n')
-----
('Con: ', 3, 'Depth: ', 1, 'http://www.colegiocastelgandolfo.cl/lista-de-utiles-2018/', '<', 'http://www.colegiocastelgandolfo.cl', u'L
e \xfatiles 2019')
-----
```

Figura 9: Mensaje desplegado por main.cpp al ejecutar el crawler.

Fuente: Elaboración Propia.

Esta funcionalidad se encarga de ejecutar el script crawler.py, archivo en python que es manejado por la librería Scrapy. La idea es que este programa en C++ gestione las múltiples llamadas al crawler.py con diferentes opciones, ya sea cambiando el nombre del output generado por Scrapy como gestionando el inicio y término de cada output según el límite de extracciones por archivo (100000 urls crawladas por archivo por default), ya que después de este límite se comienza a crawllear sitios en otro archivo con diferente nombre incremental (crawl1.csv, crawl2.csv, crawl3.csv, etc).

Después de alcanzar el límite máximo de rastreos por archivo el Crawler dormirá 20 Segundos, mostrando el siguiente mensaje por consola. Ver figura (10).

```
-----
('There are', 0, 'working links and', 0, 'broken links.')
-----
DURMIENDO 20 SEGUNDOS-----
-----
```

Figura 10: Mensaje desplegado por main.cpp luego de terminar una acción.

Fuente: Elaboración Propia.

Este momento es crucial, ya que es el único momento del crawling en el que se tiene certeza de que no hay *threads* corriendo en segundo plano y que todos han terminado normalmente, por lo que hacer una detención de ejecución con el comando CTRL+C en consola es totalmente seguro en este momento, ya que no habrá pérdida de datos

ni de links encolados que falten por crawlear, puesto que todas las variables necesarias para continuar normalmente en una nueva ejecución se guardan en la carpeta "crawls" y que es gestionada por Scrapy al momento de crawlear sitios web.

Adicionalmente, es posible detener el crawler antes de que llegue al límite de urls por archivo y es usando el comando CTRL+C en consola, pero solo una vez, ya que esto le dice a Scrapy que detenga el crawling de **manera segura**, es decir, terminando normalmente todos los threads en ejecución. Es por esto que la detención tampoco es inmediata.

La aplicación del comando CTRL+C más de una vez le dice a Scrapy que detenga el crawling de **manera forzosa**, es decir, inmediatamente, perdiendo información valiosa que se usaba en cada thread.

Después de terminar con el crawling de cada archivo, se comprueba si existen archivos CSV en la carpeta "outputs" que no han sido importados a SolR (nombres de archivos importados se guardan en imported_files.txt) y se procede a importarlos si es que se alcanza la cantidad mínima de archivos para importar. Este paso de importación también puede hacerse por separado y se explica en más detalle en los siguientes comandos.

Luego que pasan los veinte segundos de espera, el crawler seguirá el crawling en otro archivo, si es que aún quedan páginas encoladas para crawlear.

El proceso se repite hasta que el crawler duerma dos veces sin haber cambiado visited.txt, que es el archivo que guarda todos las urls crawleadas.

Ejecución del Re-Crawler o Actualizador de Índice

`$ make recrawl` (21)

Este comando ejecuta el recrawler o actualizador de índices, es decir, ejecuta el archivo Main.cpp en modo recrawler. Esta funcionalidad se encarga de ejecutar el script recrawler.py, archivo en python que es manejado por la librería Scrapy y que tiene como finalidad recorrer los links que se encuentran en el archivo visited.txt y hacer crawling de ellos para extraer el texto correspondiente a su sitio.

Esta funcionalidad está pensada para actualizar la información de los sitios que ya se rastrearon pero que se consideran demasiado antiguos, por lo que hace falta actualizar.

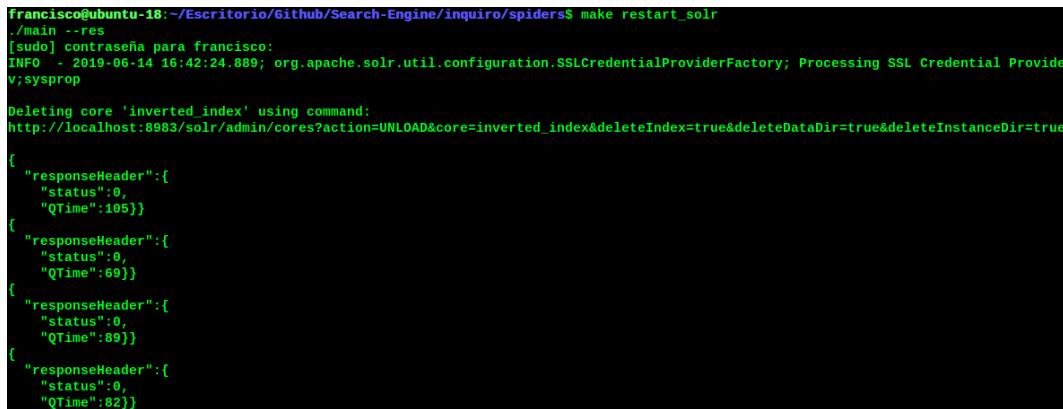
Reinicio de la Colección de SolR

`$ make restart_solr` (22)

Este comando ejecuta Main.cpp en modo reiniciar Solr y lo que hace es justamente eso. Realiza las siguientes acciones en orden.

1. Primero elimina la colección actual "inverted_index". (es decir, elimina el índice)
2. Crea la colección "inverted_index" de nuevo.
3. Se crean los Fields de Solr (body, doc_type, title, url, anchor_text, date y domain)
4. Se define al Field Url como la clave primaria de cada documento.
5. Se da permisos de escritura a la carpeta principal de Solr y se especifica ruta donde se guardará el índice de Solr.
6. Por último, se reinicia el servicio de Solr

Al ejecutarse todas las acciones exitosamente, el índice de la colección está listo para importar nuevos documentos. En la figura (11) se muestran algunos mensajes desplegados en la ejecución.



```
francisco@ubuntu-18:~/Escritorio/Github/Search-Engine/inquirol/spiders$ make restart_solr
./main --res
[sudo] contraseña para francisco:
INFO - 2019-06-14 16:42:24.889; org.apache.solr.util.configuration.SSLCredentialProviderFactory; Processing SSL Credential Provider
V;sysprop
Deleting core 'inverted_index' using command:
http://localhost:8983/solr/admin/cores?action=UNLOAD&core=inverted_index&deleteIndex=true&deleteDataDir=true&deleteInstanceDir=true
[
  "responseHeader":{
    "status":0,
    "QTime":105}}
[
  "responseHeader":{
    "status":0,
    "QTime":69}}
[
  "responseHeader":{
    "status":0,
    "QTime":89}}
[
  "responseHeader":{
    "status":0,
    "QTime":82}}
```

Figura 11: Mensaje desplegado por main.cpp luego de reiniciar la colección de Solr.

Fuente: Elaboración Propia.

Importación de archivos CSV a Solr

\$ make import_files (23)

Este comando ejecuta Main.cpp en modo importar archivos, es decir, se buscan archivos CSV no importados en la carpeta outputs y si se encuentran, se importan a la colección de Solr. En la figura (12) se muestran algunos mensajes desplegados en la importación.

Re-Importación de archivos CSV a Solr

\$ make reimport_files (24)

Este comando ejecuta Main.cpp en modo re-importar archivos, es decir, se buscan archivos CSV no importados en la carpeta reoutputs y si se encuentran, se importan a la colección de Solr, reemplazando el contenido de los documentos que tengan igual url.


```
francisco@ubuntu-18:~/Escritorio/Github/Search-Engine/inquiro/spiders$ make import_files
./main --imp
-----
IMPORTING: crawl1.csv-----
{
  "responseHeader":{
    "status":0,
    "QTime":1313}}
-----
!!!IMPORTACIÓN FINALIZADA-----
-----
EJECUCION FINALIZADA!-----
-----
```

Figura 12: Mensaje desplegado por main.cpp al importar archivos CSV.

Fuente: Elaboración Propia.

Reiniciar Colección Solr e Importar archivos CSV

\$ make restart_and_import (25)

Este comando es una combinación de los dos comandos anteriores, es decir, primero se reinicia la colección Solr y luego se importan los archivos CSV que no hayan sido importados.

Reiniciar Colección Solr y Correr Crawler principal

\$ make restart_and_run (26)

Este comando es una combinación de dos comandos anteriores, es decir, primero se reinicia la colección Solr y luego se corre el crawler normalmente. Rastreando páginas e importando en cada transición de archivos CSV.

Limpiar todos los archivos

\$ make full_clean (27)

Este comando es de los más útiles, ya que limpia totalmente el crawler de archivos residuales de una ejecución anterior. En resumen, se remueven los siguientes archivos:

1. Todos los archivos CSV de la carpeta outputs
2. Todos los checkpoints de la carpeta crawls
3. La carpeta Logs con todos los logs en su interior
4. El archivo imported_files.txt
5. El archivo reimported_files.txt

6. El archivo visited.txt
7. El archivo revisited.txt
8. Todos los archivos de compilación python

Luego de esto, el sistema se encuentra exactamente en el mismo estado inicial, antes de cualquier comando o ejecución. En otras palabras, se considera lo que se llama muchas veces *hard reset*.

- **Verificación de carpetas** Esta funcionalidad es una de las tareas del controlador más simples pero a la vez más importantes, ya que para que todas las demás funcionalidades tengan efecto, todas las carpetas que almacenan partes importantes de los resultados del crawler deben estar creadas y listas para su funcionamiento. En resumen, las carpetas más importantes son las siguientes.

- **Logs:** Carpeta ubicada en el directorio "spiders", almacena los logs principales del sistema: main_log.txt y domains_log.txt, de los cuales se hablará más en detalle en las siguientes secciones.
- **Outputs:** Carpeta también ubicada en "spiders", y como su nombre lo dice, almacena los *outputs* principales del crawling, es decir, todo el texto extraído de los sitios web que es almacenado en archivos CSV.
- **Reoutputs:** Carpeta muy similar a la anterior, también ubicada en "spiders", y se encarga de almacenar los outputs de un proceso de "re-crawling", es decir, una actualización de índice de sitios ya rastreados anteriormente.
- **Data:** Carpeta ubicada en el directorio "solr", es de las más importantes, ya que se encarga de almacenar los datos del índice de SolR.

- **Configuración Logging**

Como se explicó en secciones anteriores, la funcionalidad de *logging* en sistemas complejos como este es esencial, ya que permite ver y analizar el comportamiento del sistema en todo momento. Esto incluye informaciones útiles para los administradores, como los horarios y el tipo de acciones que se ejecutan.

En el componente "Crawler.py" anteriormente explicado se menciona un sistema de logging interno el cual es gestionado por *Scrapy*, pero esta configuración es diferente, ya que se trata de un sistema implementado en C++ a más bajo nivel.

Más adelante se darán más detalles sobre la implementación y funcionalidad sobre esto.

El controlador posee un rol importante en este sentido, debido a que es el responsable de realizar las configuraciones pertinentes para que esta funcionalidad sea lo más útil posible. Entre todas las configuraciones, las principales son las siguientes.

- **Patrón de Log:** El formato en que la información *logging* es presentada por defecto no siempre es la mejor, por lo que definir un patrón propio del sistema a

analizar es indispensable para lograr un entendimiento claro y efectivo. En este caso, el formato utilizado es el siguiente.

$$[%d / \%m / \%Y - \%T][\%l] \%v \quad (28)$$

Que utiliza la siguiente simbología definida por la librería Spdlog.

- *%d*: Día del mes. Se representa como un entero de dos dígitos, con un rango que va desde 01 a 31.
 - *%m*: Mes del año. Se representa como un entero de dos dígitos, con un rango que va desde 01 a 12.
 - *%Y*: Año actual. Se representa como un entero de cuatro dígitos.
 - *%T*: Tiempo actual, en Horas, minutos y segundos. Cada uno de estos números se representa como un entero de dos dígitos. Las horas con un rango de 00 a 23 y los minutos y segundos con un rango de 00 a 59.
 - *%l*: Tipo de Log. Se definen como los tipos de información que entrega cada Log. Por defecto se definen cuatro categorías en inglés: *info*, *warning*, *error* y *debug*, que representan informaciones, alertas, errores y mensajes de revisión respectivamente.
 - *%v*: Descripción del Log. Breve texto que explica la acción particular de cada mensaje Log.
- **Archivo Log por defecto**: Se indica brevemente cuál será el nombre del archivo que contendrá todos los mensajes *log* y cuál será su ubicación.

3.5.2. Crawler.py

Como se explicó anteriormente, el componente *Crawler.py* es de los mas importantes, ya que es el responsable directo de la extracción de información de sitios web. También se dieron detalles de su funcionamiento, de las entradas y de sus salidas. Ahora es interesante hablar de esto desde una mirada más global.

El controlador es el responsable de ejecutar las acciones que se indican en los comandos *Makefile*, y una de estas acciones tiene directa relación con este componente, que es la de ejecutar el crawler, comando **make run** en la tabla (4).

Luego de que todo el procesamiento de *Crawler.py* se realiza, los *outputs* CSV que contienen todo el texto rastreado son gestionados por el Controlador para controlar la importación de estos en la colección principal. En el componente *SolR* se darán mas detalles de esto.

A continuación se explican las funcionalidades principales de este componente, vistos de una perspectiva más global.

- **Extracción páginas Web**: Claramente es la funcionalidad principal de todo *crawler* o rastreador, pero este lo hace de una forma bien particular. Se trata de un script python

que junto a la librería *Scrapy* se encarga de extraer, parsear y almacenar todas las palabras que se encuentran en los sitios web del dominio chileno en archivos CSV, para que luego estos archivos puedan ser usados para proveer de términos al índice invertido de SolR. El cual es el motor de búsqueda principal usado en este proyecto para la recuperación eficiente de documentos.

- **Actualizador de Índices:** Si bien, esta funcionalidad le corresponde al *spider* *recrawler.py*, se considera dentro del mismo componente debido a que son técnicamente iguales. Anteriormente se había explicado que la única diferencia entre *crawler.py* y *recrawler.py* era que el segundo no posee la capacidad de seguir rastreando enlaces de páginas que se encuentren en el sitio actual, sino que sólo rastrea los sitios que se encuentren en *visited.txt*.

De igual forma, el progreso del proceso de actualización también se va guardando en otro archivo de texto plano llamado *revisited.txt*, de forma que se puedan hacer pausas y reanudaciones del proceso de la misma manera en que lo hace el rastreador principal.

Esta funcionalidad está pensada para actualizar la información de los sitios que ya se rastrearon pero que se consideran demasiado antiguos, por lo que hace falta actualizar su información.

- **Generación de Grafo:** Esta funcionalidad se considera un proceso extra, pero que puede ser muy valioso en el futuro para otras investigaciones. Hablamos de la generación de un grafo que representa la web chilena actual. De forma que cada sitio visitado sea un nodo y cada enlace entre estos sea una unión en el grafo. La representación implementada se trata de un archivo de texto plano llamado *graph.txt*, muy similar a *visited.txt*, pero que almacena cada url de una páginas visitada, y junto a ella la url de una páginas a la que apunta, es decir, cada línea de este archivo contiene dos url separadas por un espacio, que representan la unión.

- **Logging Interno:** Hay que tener claro que esta funcionalidad *Logging* es interna, es decir, es propia de este componente y que gracias a *Scrapy* es posible modificar el formato de mensajes de forma conveniente y almacenarlo en el archivo *domains_log.txt*, dentro del directorio *Logs*.

En este caso, el logging interno es usado principalmente para el seguimiento de los dominios rastreados, de forma que se sepa en todo momento cual es el dominio actual al cual se está extrayendo información y saber en qué fecha y hora ocurren cambios de dominio.

3.5.3. Solr

A pesar de que también hemos mencionado reiteradas veces a SolR, este sub-componente es algo más particular, ya que no se trata directamente de la *API de Apache*, sino que más bien del módulo dentro de *Main.cpp* que se encarga de comunicarse con el servidor de SolR.

Esto incluye tanto configuraciones como acciones que tengan relación con la colección del índice invertido de *SolR*, y que se procede a explicar a continuación.

- **Configuración Esquema:** Como toda colección de índices, esta debe tener un formato o esquema que define a los datos como tal, para tratar cada información o atributo de la forma más eficiente posible según su tipo de dato.

En *SolR* la indexación funciona almacenando documentos dentro de colecciones, y estos a su vez almacenando atributos que describen a cada documento. Estos atributos pueden ser de muchos tipos que ya están previamente definidos, pero la diferencia está en que ciertos tipos de datos poseen propiedades o funcionalidades adicionales, lo que implica que una buena elección de tipos de datos para cada atributo debe ser tomada adecuadamente.

En este caso, se escogen variados tipos para cada uno de los campos a guardar por sitio y que se resumen en la tabla (5).

| Campo | Tipo de dato | Propiedad |
|----------|--------------|---|
| url | string | Identificación primaria |
| body | text_es | Tokenizador, LowerCase, Stopwords, Spanish Stemming |
| doc_type | text_es | Tokenizador, LowerCase, Stopwords, Spanish Stemming |
| title | text_es | Tokenizador, LowerCase, Stopwords, Spanish Stemming |
| date | text_es | Tokenizador, LowerCase, Stopwords, Spanish Stemming |
| domain | text_es | Tokenizador, LowerCase, Stopwords, Spanish Stemming |
| depth | plongs | Apoyo Estadísticas |
| con | plongs | Apoyo Estadísticas |

Tabla 5: Campos de la colección de *SolR*.

Fuente: Elaboración Propia.

Como se puede observar, la *url* posee un tipo de dato "*string*" en la colección, y eso se debe a que este campo es la clave primaria de cada documento, es decir, identifica a cada uno de estos y los hace únicos. De forma que si se importa un nuevo documento con la misma *url* este último reemplazará el anterior y se entenderá que se trata de una actualización.

Ahora bien, con respecto al *body*, título, texto ancla, tipo de documento y fecha, se decide utilizar el tipo de dato *text_es*, el cual hace referencia a un texto en español que posee las siguientes propiedades.

- **Tokenizador:** Propiedad que permite separar oraciones en tokens individuales, es decir, se trata a cada palabra de la oración como un solo término, lo que permite buscar documentos que los contengan adecuadamente.
- **Lowercase:** Esta propiedad permite transformar cada término indexado en una palabra que solo contenga letras minúsculas, lo que facilita la recuperación de documentos relevantes para consultas similares que contengan letras mayúscula.

- **Stopwords Removal:** Propiedad fundamental para un motor de búsqueda, ya que permite eliminar palabras sin significado, llamadas stopwords, de forma que solo se almacenen términos con significado y que sean relevantes para la recuperación de documentos.
- **Spanish Stemming:** Esta es otra de las propiedades fundamentales de un motor de búsqueda, ya que permite almacenar solamente la raíz sintáctica de cada término, de forma que cada vez que se haga una consulta con un término perteneciente a la familia de este término, se tenga en cuenta al momento de evaluar la relevancia de los documentos.

Básicamente todos los campos que contengan más de una palabra y que contengan términos relevantes para las búsquedas, deben tener este tipo de dato.

Por último, tenemos los últimos dos campos: *con* y *depth*, que poseen el tipo de dato *plong*, que corresponde a un valor numérico de gran tamaño.

Este tipo de dato es útil especialmente al momento de hacer estadísticas, ya que permite realizar análisis y filtros rápidos según los valores numéricos que se posean para cada documento.

■ Reinicio de Colecciones

El reinicio de colecciones también es una funcionalidad gestionada a través del Controlador, que recibe el comando para reiniciar y luego este es ejecutado internamente por *main.cpp* para mandar señales al server de SolR, para que ejecute acciones de reinicio.

Las tareas más importantes que se realizan en esta funcionalidad son las siguientes.

1. **Eliminar colección "inverted_index" actual:** Para poder hacer un reinicio efectivo de la colección de SolR hace falta eliminar todo el índice, ya que esto asegura que no hayan documentos residuales de índices anteriores.
2. **Crear colección "inverted_index" nueva:** Luego de eliminar la colección, es necesario crear un índice nuevo, para poder importar documentos que ya han sido rastreados previamente.
3. **Crear campos de SolR:** Se crean los *Fields* o campos de Solr: cuerpo, tipo de documento, título, url, texto ancla, fecha y dominio, con los tipos de datos que se explicaron anteriormente y que permite tratar cada tipo de información de una manera eficaz.
4. **Asignación de clave primaria:** Para poder identificar a cada documento de forma única se define al campo url como la clave primaria de cada documento, ya que no es posible tener dos documentos distintos con igual url en la web.
5. **Actualización de permisos:** Para poder guardar y editar el índice de SolR, se otorgan permisos de escritura a la carpeta principal de Solr y se especifica la ruta donde se guardará el índice de SolR. Este es un paso de gran importancia, ya que

muchas veces la partición raíz de los sistemas *linux* no poseen tanto espacio como lo son las particiones de usuario, por lo que usar ubicaciones con mas espacio significa poder gestionar índices más grandes en espacio.

6. **Reiniciar servicio SolR:** Por último, se termina el proceso reiniciando el servicio cliente-Servidor de SolR, es decir, se vuelve a iniciar el servidor que aloja a SolR, actualizando sus configuraciones.

Luego que se ejecutan todas estas acciones exitosamente, la colección queda lista para indexar nuevos documentos.

■ Importación de Colecciones

La importación también es una acción desencadenada por la gestión de los comandos Makefile del Controlador, ya que, en un principio, es el usuario el que solicita la importación directa de documentos.

En la tabla (4), el comando para importar documentos no importados desde la carpeta *outputs* es **make import_files**.

Para evitar la re-importación de documentos en esta carpeta, se gestiona una lista de nombres de archivos importados, que se actualiza cada vez que un nuevo archivo es importado exitosamente. Esta lista de nombres es guardada en un fichero de texto llamado *imported_files.txt*, y es analizada cada vez que se realiza una nueva importación, de forma que cada vez se verifica cual de los archivos actuales no están presentes en la lista y que por lo tanto no han sido importados.

Ahora bien, hasta el momento hemos hablado solo de la importación explícita de documentos, pero esta importación también se realiza automáticamente al momento de rastrear nuevos sitios web, ya que cada vez que se alcanza el número máximo de urls por archivo de crawling, el proceso es detenido y nuevamente se verifica si hay archivos CSV sin importar para agregarlos a la colección, pero esto solo puede suceder si se cumplen las siguientes dos condiciones.

- **Si existen suficientes archivos para importar:** En el script *crawler.py* se puede especificar una constante numérica que corresponde al número mínimo de archivos CSV no importados que tienen que existir en la carpeta *outputs* para que se realice una importación, es decir, si se supera este número, la importación puede seguir su curso.
- **Si la importación se realiza en horario nocturno:** Además de poseer el número mínimo de archivos no importados, es necesario que el momento de la verificación de archivos se realice en horario nocturno, que por defecto entre las 2 am y las 6 am. Esto garantiza el mayor tiempo de rastreo de sitios posible y además garantiza la disponibilidad online del buscador la mayor parte del tiempo en que las personas de Chile están activas.

Al cumplirse estas dos condiciones, la importación automática puede llevarse a cabo cada vez que el *crawler* vaya avanzando con la extracción de información.

3.5.4. Spdlog

Spdlog no se considera un componente como tal, sino más bien una librería externa que ayuda al controlador a ejecutar funcionalidades de *logging* y que poseen las siguientes características.

■ Mensajes Multipropósito

Los tipos de mensajes que ofrece esta librería poseen una gran variedad, y es que se pueden usar prácticamente para todo lo que se necesite. Por defecto, existen cuatro tipos de *logs* posibles y que son los siguientes.

- **Información:** Mensajes que no tienen otra finalidad más aparte de mostrar o anunciar una descripción de lo que está sucediendo al momento del despliegue de este.
- **Alerta:** Como su nombre lo dice, son mensajes especiales que tienen la función de alertar algo más relevante que una simple información. Generalmente se usa en secciones críticas del sistema como el inicio o término de procesos relevantes.
- **Error:** También como lo dice el nombre, se trata de mensajes que solo se despliegan cuando un proceso no sigue su curso normal de funcionamiento debido a alguna razón. Para el uso de estos mensajes, se deben prever los posibles errores que pueden suceder al momento de ejecutar tareas críticas.
- **Depuración:** Los mensajes de depuración ayudan al proceso de identificar y corregir posibles errores de programación. Son utilizados generalmente cuando se necesita encontrar la causa de un error.

■ Patrón de Logging:

Como toda librería de Logging, esta tiene un patrón configurable y ajustable a las necesidades de cada sistema.// Por defecto, el patrón que sigue esta librería es el siguiente.

$[YYYY - MM - dd/HH : mm : ii] [Nombre Log] [tipo Log] Descripción$ (29)

Siendo *YYYY*, *MM* y *dd*, el año, el mes y el día en que se crea el mensaje *logging* respectivamente. También *HH*, *mm* y *ii* son la hora, minuto y segundo, que representan la hora exacta de creación del mensaje respectivamente.

Por otro lado, también se tiene la descripción del mensaje, el nombre del *logging* como tal y el tipo de este, los cuales son los mismos definidos anteriormente en la sección de mensajes multi-propósito.

Pero este patrón no es tan simple de seguir para el sistema actual, ya que para empezar, el formato de fecha no es el que se usa en Chile, por lo que es conveniente cambiarlo. También el nombre de *logging* no es tan necesario, ya que solo existe uno en este

módulo, por lo que sería una información irrelevante y que se puede omitir. Teniendo en cuenta estas observaciones, finalmente se utiliza el siguiente patrón.

$$[dd - MM - YYYY/HH : mm : ii] [tipo Log] Descripción \quad (30)$$

Que posee igual nomenclatura, pero en diferente orden, más simple y omitiendo información innecesaria.

Se ha hablado bastante de *logging*, pero solo porque es importante recalcar su finalidad e importancia en el sistema.

3.6. Diseño de Interfaz

La Interfaz Gráfica de Usuario (GUI por su nombre en inglés, Graphical User Interface) es parte fundamental de cualquier aplicación. Al comenzar a trabajar con una computadora el usuario comienza a interactuar con la Interfaz, ya sea la del sistema operativo, la de un software en particular o la de cualquier sitio web. Es donde comienza la interacción hombre-computadora. En ocasiones, a partir de la GUI se puede determinar si una aplicación será utilizada o no para resolver los problemas para los cuales fue diseñada. En el actual mundo digital se debe ofrecer al usuario una Interfaz que lo ayude a concretar las tareas de manera rápida, sencilla y satisfactoria. Es la Interfaz la responsable de ofrecer una interacción fluida y agradable [Albornoz et al., 2017].

Por estas razones, el nuevo buscador *Inquiro* no puede quedarse atrás y debe tener una interfaz que corresponda a los tiempos actuales.

Es por esto que se diseña una interfaz responsiva, es decir, que sea capaz de adaptarse a diferentes dispositivos y resoluciones.

Cada interfaz de escritorio está pensada para computadores con pantallas estándar, es decir, de más de 768 píxeles de resolución. En cambio las versiones móviles, son para pantallas de menos de 768 píxeles.

A continuación, se presentan los diseños iniciales del sistema, tanto en su versión de escritorio como en su versión móvil.

3.6.1. Vista principal de Escritorio

Vista principal de presentación del buscador, en donde la principal funcionalidad es sugerir el ingreso de términos para la búsqueda.

InquiRo Chile

El buscador de contenidos de Chile

Figura 13: Interfaz gráfica de la páginas web principal de *InquiRo.cl* .
Fuente: Elaboración Propia.

3.6.2. Vista principal Móvil

Vista principal de presentación del buscador, pero en su versión móvil, posee la misma funcionalidad de su versión de escritorio, que es sugerir el ingreso de términos para la búsqueda.

InquiRo Chile

El buscador de contenidos de Chile

Figura 14: Interfaz gráfica de la páginas web móvil principal de *InquiRo.cl* .
Fuente: Elaboración Propia.

3.6.3. Vista de Resultados de Escritorio

Corresponde a la vista de escritorio de los resultados de una búsqueda previa. Los documentos más relevantes aparecen más arriba en la lista de sitios entregados.

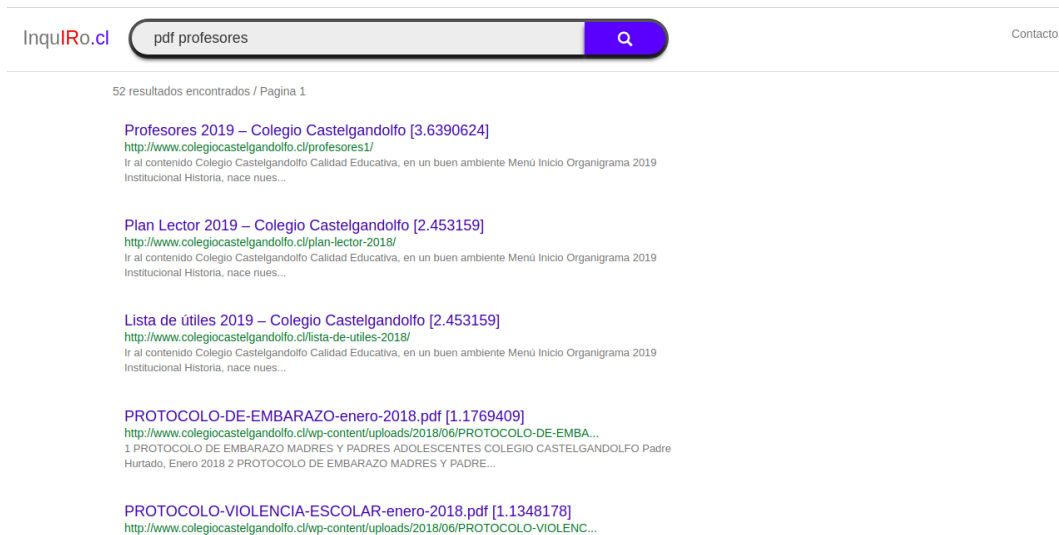


Figura 15: Interfaz gráfica de la páginas web de Resultados de *InquiRo.cl* .
Fuente: Elaboración Propia.

3.6.4. Vista de Resultados Móvil

Corresponde a la versión móvil de la vista de resultados de una búsqueda. De igual forma, los documentos más relevantes aparecen más arriba en la lista de sitios entregados.

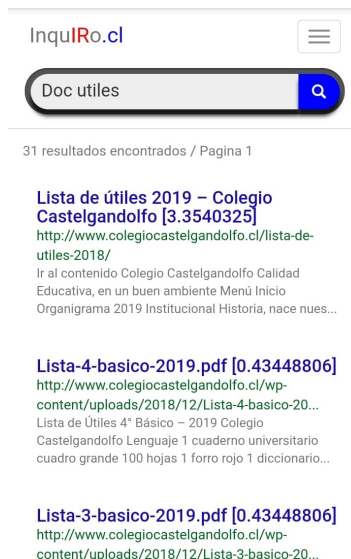


Figura 16: Interfaz gráfica de la páginas web móvil de Resultados de *InquiRo.cl* .
Fuente: Elaboración Propia.

Como se puede observar en las interfaces presentadas anteriormente, estas poseen una gran variedad de colores, pero estos no fueron escogidos al azar, ya que tienen una finalidad bien específica que se procede a explicar por conjunto de colores.

- **Blanco, azul y rojo:** Representan los colores de la bandera de Chile. Como se explicó en el capítulo uno, este proyecto está inspirado en un antiguo buscador llamado *Inquiro* también, pero que poseía la habilidad de buscar en colecciones de otros países como Argentina, México, entre otros. Por lo que es importante resaltar que este nuevo proyecto con el mismo nombre solo está pensado para Chile.
- **Azul, verde y gris:** Estos colores son usados principalmente en la vista de resultados, ya que a lo largo de los años se ha visto un gran uso de estos en buscadores populares como *Google*, *Yahoo* o *Bing*, los cuales han experimentado con pruebas de usabilidad más de una vez, y que por lo tanto, demuestran ser los colores más prácticos a la hora de encontrar recursos web en motores de búsqueda.

3.7. Implementación

Después de ver todo el diseño explicado en las secciones anteriores, hay que seguir con la manera en que se implementa la solución, ya que esto muchas veces define el comportamiento del sistema y su rendimiento.

A continuación se definen todos los recursos de *software* y *hardware* utilizados, además de hablar sobre aspectos generales de la implementación como tal.

3.7.1. Recursos utilizados

En cuanto a los recursos que se utilizaron para la implementación de *Inquiro*, se puede dividir en dos grandes categorías, los recursos de *software* y los recursos de *hardware*. Los primeros se definen como un conjunto de programas y rutinas que permiten a las computadoras realizar determinadas tareas. El segundo se refiere a un conjunto de elementos físicos o materiales que componen una computadora o un sistema informático.

Recursos de Software

En la tabla (6) se resumen las principales herramientas de *software* utilizadas.

El primer recurso a revisar es *Sublime Text 3*, corresponde a un poderoso IDE que permite programar y editar texto fácilmente, resaltando las palabras clave de cada lenguaje y remarcando los cambios efectuados desde el último *commit* del control de versiones.

El segundo es *SolR*, ya se ha hablado bastante de esta herramienta a lo largo de esta memoria, pero se trata del motor de búsqueda principal del proyecto, ya que provee un poderoso sistema de búsqueda e indexación que facilita la implementación web del buscador. Esto se logra a través de peticiones HTTP al servidor local de *SolR*, ya que posee una arquitectura cliente-servidor que facilita la recuperación de documentos desde el índice invertido hacia el sitio web.

| Herramienta | Versión | Utilidad |
|----------------|---------|--|
| Sublime Text 3 | 3.2.1 | Herramienta utilizada como editor de texto y código fuente para el desarrollo del proyecto |
| SolR | 8.0.0 | Herramienta utilizada como motor de búsqueda |
| Django | 2.1.7 | Framework python utilizado para el diseño web del buscador |
| Gitkraken | 6.0.0 | Interfaz gráfica git utilizada para el control de versiones del proyecto |
| Ubuntu | 18.04.2 | Sistema Operativo utilizado para el proyecto. |

Tabla 6: Recursos de Software utilizados en el desarrollo de *Inquiro.cl*

Fuente: Elaboración Propia.

Como tercera herramienta tenemos a *Django*, la cual se define como un framework de desarrollo web basado en el lenguaje Python. Ya se ha hablado bastante de esta herramienta en los capítulos anteriores, pero es importante recalcar la importancia que posee.

Este *framework* provee módulos simples para diseñar las interfaces de usuario y la lógica del buscador, y esto lo hace a través del uso de *templates* y controladores respectivamente.

Los *templates* permiten definir código HTML común, pero con una particularidad, lo hacen a través del uso de código *python* que ayuda a gestionar estas etiquetas HTML dinámicamente, es decir, cada objeto o entidad que es gestionada en el Controlador de este *framework* puede ser pasada a la vista para que se despliegue de una manera simple y eficiente. Por otro lado, los controladores se encargan de definir toda la lógica del sitio. En este caso, contienen toda la implementación del buscador, que evidentemente posee la función principal de buscar documentos relevantes para ciertas consultas.

Gitkraken por su parte, es una herramienta que permite un manejo gráfico y simple del control de versiones de *git*. Gracias a esta interfaz, es más rápida la implementación de cambios en el proyecto, ya sean cambios en el *crawler*, en el índice de *SolR* o incluso en la páginas web del buscador.

Por último, tenemos el sistema operativo *Ubuntu*, que fue utilizado para la memoria. Se utilizó en prácticamente todo el proyecto, y esto se debe a que posee una gran compatibilidad con muchas herramientas de desarrollo, incluidas las antes mencionadas.

En otros sistemas operativos los módulos como *Scrapy* no funcionan de la misma forma y pueden incluso a generar errores de compatibilidad.

Recursos de Hardware

En la tabla (7) se resumen las principales herramientas de *hardware* utilizadas.

El servidor de *Inquiro* es parte de la red de servidores que posee la Universidad Técnica Federico Santa María (UTFSM), y este paso por un proceso de limpieza antes de alojar los ser-

| Tipo | Características |
|----------|--|
| Servidor | Intel Xeon E31220 3.10GHz, 4GB Ram, 1HDD 220GB |
| Notebook | Intel Core i7-8550U 1.80GHz, 8GB Ram, 1 HDD 1TB, 1 SSD 128GB |

Tabla 7: Recursos de Hardware de Inquiro.

Fuente: Elaboración Propia.

vicios de esta memoria. Básicamente se eliminaron los archivos innecesarios y se instalaron los módulos de *Inquiro* a través de comandos *git* que descargan el contenido del proyecto fácilmente.

Además de este servidor, también se utilizó un notebook en el que se produjo todo el proceso de desarrollo, ya sea creando nuevas funcionalidades, haciendo pruebas de crawling con sitios pequeños, probando el despliegue del sitio web del buscador o controlando el servidor remotamente a través de comandos SSH.

3.7.2. Implementación General

Como se mencionó en la sección anterior, la implementación del sistema diseñado en el servidor se basa en la descarga del repositorio *git* del proyecto en el directorio deseado.

Al descargar el repositorio, se descargan los módulos principales que utiliza el sistema, pero para funcionar este debe contar con los siguientes paquetes previamente instalados:

- Compiladores C++, Java y Python
- Servicio SolR
- Pip: Gestor de Paquetes python
- Scrapy
- Tika
- Virtual Enviroment Python3
- Nginx

Además, se sugiere la instalación del paquete *Screen* para *Linux*, ya que permite gestionar sesiones de terminales en el servidor. De esta forma se pueden ejecutar tanto el *crawler* como el despliegue web simultáneamente en distintas sesiones.

Una vez que todos estos requerimientos son instalados satisfactoriamente, se deben hacer algunas pequeñas configuraciones para que tanto el *crawler* como el sitio web funcionen adecuadamente y que se explican a continuación.

Implementación Crawler

Lo primero que se debe hacer para la implementación del *crawler* es verificar que todo funcione correctamente. Para esto basta con revisar que todas las opciones de este módulo funcione correctamente.

En el directorio *Spiders* dentro de *Inquiro*, se encuentra un fichero *Makefile* que controla todo el sistema con los comandos vistos en (4).

Se sugiere primero ingresar el comando *make build* para compilar el *script* principal del sistema llamado *main.cpp* y así poder utilizar los demás comandos disponibles.

Una vez que el *script* ya haya compilado satisfactoriamente, se puede ejecutar el comando *make run* para verificar que *Scrapy* y *Tika* se hayan instalado correctamente.

Por defecto, el sistema usa sitios semilla de prueba para justamente verificar lo mencionado anteriormente, pero una vez que se quiera ejecutar el *crawler* de manera definitiva, debe cambiarse esta opción a las semillas definitivas que se encuentran en *cl_seeds.txt*.

Al mismo tiempo, debe verificarse que el servicio de *SolR* este siendo ejecutado en segundo plano en el puerto 8983, de forma que la importación de archivos hacia el índice se realice de manera normal.

Por defecto, el *crawler* importa archivos al índice invertido automáticamente luego de terminar con el rastreo en un archivo y que al mismo tiempo exista una cantidad mínima de archivos no importados en la carpeta *outputs*. Aunque todo esto es válido si sucede en la madrugada de cualquier día, ya que de lo contrario el *crawler* continuará rastreando sitios en un archivo CSV nuevo.

Finalmente, luego de que todas las pruebas hayan sido satisfactorias y todas las configuraciones sean correctas, se puede ejecutar el comando *make full_clean* para eliminar todos los archivos residuales de una ejecución anterior, de forma que se pueda empezar la ejecución definitiva desde cero.

Implementación Sitio Web

Para implementar el Sitio Web, primero que todo hay que tener instalado los paquetes anteriormente mencionados, especialmente *pip* y el *virtual enviroment de python3*, ya que esto permite la instalación automática de las librerías que se usan en el sitio web.

Primero que todo, en el directorio *webpage* se debe crear un entorno virtual de *python3*, activarlo y luego instalar las librerías necesarias con *pip* y el fichero *requirements.txt*.

Una vez instaladas estas librerías, es necesaria la configuración de *Nginx* para permitir el despliegue de sitios web en la máquina actual. Además, se debe configurar el fichero *webpage_uwsgi.ini* con los parámetros correspondientes a la máquina actual, para que se pueda enlazar el sitio del buscador con el servicio de *Nginx*.

Una vez que el servicio de *Nginx* se está ejecutando en segundo plano y todas las configuraciones han sido realizadas, puede activarse el entorno virtual y utilizar la librería *Uwsgi* para ejecutar el fichero `webpage_uwsgi.ini` y desplegar el sitio web.

Finalmente, si se enlaza el dominio www.inquiro.cl de NIC Chile¹⁷ con el servidor configurado, se puede visualizar el buscador funcionando a la perfección.

¹⁷www.nic.cl

CAPÍTULO 4

VALIDACIÓN DE LA SOLUCIÓN

En este capítulo se presenta la validación de la solución propuesta, es decir, se comprueba que el trabajo realizado en los capítulos anteriores es válido para solucionar o mitigar el problema central de esta memoria.

A continuación se presentan dos secciones de validación. En la primera validación se realiza un estudio comparativo de rendimiento y resultados entre el buscador *Inquiro* y los buscadores similares más populares actualmente.

En la segunda validación se realizan estadísticas de la colección construida con el *crawler*, con la finalidad de comprobar los objetivos principales planteados en esta memoria.

4.1. Comparación de Inquiro con Buscadores Actuales

En esta sección se realiza una comparación entre el buscador construido en esta memoria y los buscadores similares que más se utilizan actualmente. Mas específicamente hablando, la comparación se realiza con Google.cl¹⁸, Yahoo Chile¹⁹ y Bing Chile²⁰.

A pesar de que no todos estos buscadores poseen una versión ".cl", todos poseen opciones que permiten restringir los resultados solo a sitios chilenos, lo que permite una comparación efectiva con *Inquiro.cl*.

Esta comparación se realiza analizando los 20 primeros resultados obtenidos de 6 consultas diseñadas aleatoriamente en estos cuatro buscadores. Se comparan métricas como el número de resultados obtenidos, los tiempos de respuesta y también un análisis cualitativo de estos resultados.

Según [Broder, 2002] la taxonomía de las búsquedas *web* se puede clasificar en tres tipos fundamentales: **informacionales**, **navegacionales** y **transaccionales**. Todas estas intentan explicar la "necesidad" o intención detrás de cada consulta. A continuación se explica más en detalle cada tipo de consulta.

- **Informacionales:** La intención es adquirir cierta información que se supone que está presente en una o más páginas web de forma estática. No se necesita ninguna otra interacción, excepto la lectura. En otras palabras, que sea estática quiere decir que el documento de destino no se crea en respuesta a la consulta del usuario.

¹⁸www.google.cl

¹⁹www.yahoo.com

²⁰www.bing.com

- **Navegacionales:** La intención principal es llegar a un sitio en particular que el usuario tiene en mente, ya sea porque lo visitó en el pasado o porque asume que tal sitio existe.
- **Transaccionales:** La intención es realizar alguna actividad o transacción a través de la web. Las principales categorías para tales consultas son compras, búsqueda de servicios en la web, descarga de varios tipos de archivos (imágenes, canciones, etc.), acceso a ciertas bases de datos, localización de servidores, entre muchas otras.

Teniendo esto en cuenta, se procede a diseñar seis consultas aleatorias para la comparación de estos buscadores. Se escogen dos consultas de cada tipo, es decir, dos consultas informacionales, dos navegacionales y dos transaccionales.

Las consultas a evaluar son las siguientes:

| ID | Tipo | Consulta |
|----|---------------|------------------------|
| 1 | Informacional | profesores usm |
| 2 | Informacional | eclipse solar 2019 |
| 3 | Navegacional | radio carolina |
| 4 | Navegacional | red transantiago |
| 5 | Transaccional | comprar caliente camas |
| 6 | Transaccional | 13c vivo online |

Tabla 8: Consultas aleatorias diseñadas para la comparación de los buscadores.

Fuente: Elaboración Propia.

En la tabla (8) se pueden observar consultas comunes para cada tipo de consultas. Estas fueron inspiradas por consultas reales que realizaron recientemente familiares y amigos del autor de esta memoria. De esta forma se puede evaluar de mejor manera como se comportaría el buscador diseñado al recibir consultas reales.

En la tabla (9) se presentan el número de resultados y el tiempo de respuesta de cada una de estas consultas para cada uno de los buscadores a comparar: *Inquiro*, *Google*, *Yahoo* y *Bing*.

En esta tabla se pueden observar varios comportamientos característicos de cada motor de búsqueda. Y que se irán comentando a medida que se encuentren en cada consulta estudiada.

Se puede visualizar a simple vista que el número de resultados varía bastante dependiendo de la naturaleza de la consulta, pero no ocurre lo mismo con el tiempo de respuesta, ya que la gran mayoría da una respuesta en un tiempo de a lo más 1 segundo. Lo que deja en evidencia la eficiencia de los algoritmos de recuperación de documentos utilizados.

A continuación, los datos de la tabla (9), se grafican y estudian en detalle según el número de resultados y el tiempo de respuesta obtenidos en la práctica para cada tipo de consulta.

| ID | Consulta | buscador | Nº Resultados | Tiempo Respuesta |
|----|------------------------|----------|---------------|------------------|
| 1 | profesores usm | Inquiro | 86.119 | 0.55s |
| 1 | profesores usm | Google | 29.400 | 0.58s |
| 1 | profesores usm | Yahoo | 265.000 | 0.90s |
| 1 | profesores usm | Bing | 196.000 | 0.55s |
| 2 | eclipse solar 2019 | Inquiro | 8.888.529 | 0.70s |
| 2 | eclipse solar 2019 | Google | 1.590.000 | 0.82s |
| 2 | eclipse solar 2019 | Yahoo | 49.200.000 | 1.31s |
| 2 | eclipse solar 2019 | Bing | 20.500.000 | 0.83s |
| 3 | radio carolina | Inquiro | 236.232 | 0.75s |
| 3 | radio carolina | Google | 922.000 | 0.78s |
| 3 | radio carolina | Yahoo | 992.000 | 1.02s |
| 3 | radio carolina | Bing | 47.300.000 | 0.65s |
| 4 | red transantiago | Inquiro | 8.284.667 | 0.65s |
| 4 | red transantiago | Google | 178.000 | 0.69s |
| 4 | red transantiago | Yahoo | 186.000 | 1.12s |
| 4 | red transantiago | Bing | 144.000 | 0.71s |
| 5 | comprar caliente camas | Inquiro | 8.623.309 | 0.59s |
| 5 | comprar caliente camas | Google | 485.000 | 0.73s |
| 5 | comprar caliente camas | Yahoo | 174.000 | 1.22s |
| 5 | comprar caliente camas | Bing | 111.000 | 0.71s |
| 6 | 13c vivo online | Inquiro | 661.350 | 0.63s |
| 6 | 13c vivo online | Google | 25.500 | 0.68s |
| 6 | 13c vivo online | Yahoo | 53.600.000 | 0.85s |
| 6 | 13c vivo online | Bing | 88.100 | 0.74s |

Tabla 9: Comparación de resultados en distintos buscadores.

Fuente: Elaboración Propia.

4.1.1. Comparación de consultas informacionales

En esta sección se intenta comparar el comportamiento de los motores de búsqueda al ingresar dos consultas del tipo informacional: *profesores usm* y *eclipse solar 2019*. En la primera consulta la intención de búsqueda es obtener información sobre los profesores de la *Universidad Federico Santa María*, y en la segunda la intención es obtener información sobre el eclipse solar del presente año.

Las métricas de comparación utilizadas son el número de resultados que retornan y el tiempo que demoran en hacerlo, además de realizar un análisis cualitativo general de resultados.

Comparación en base al número de Resultados

Para empezar, en la primera y segunda consulta, se observa que en cuanto al número de resultados, *Yahoo* es el que más cantidad entrega, con un valor de 265.000 y 49.200.000

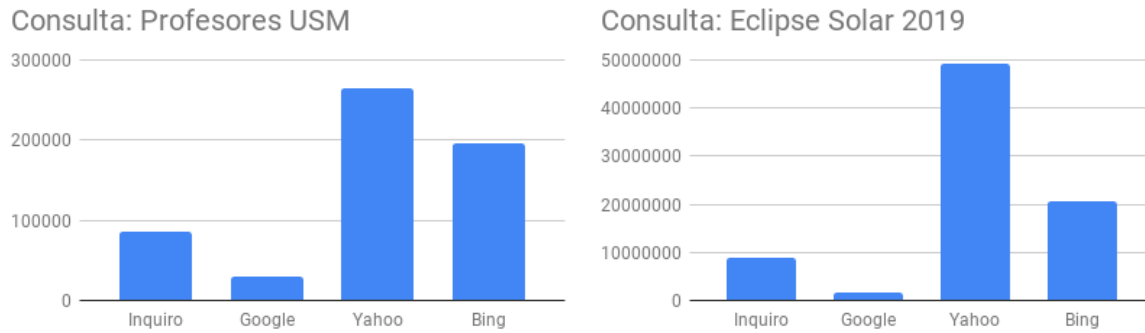


Figura 17: Comparación del N° de resultados de dos consultas informacionales

documentos respectivamente.

A diferencia de ver los resultados en la tabla, en la figura (17), se puede visualizar un patrón en común en ambas consultas. La primera con una cantidad de resultados significativamente menor que la segunda, pero que es suficiente para dejar en evidencia varias cosas.

Google a pesar de ser el buscador por excelencia y poseer la mayor cantidad de documentos indexados en el mundo, en este tipo de consultas informacionales tiende a retornar el menor número de documentos de los cuatro buscadores. Lo cual no es necesariamente desfavorable, porque puede que los documentos retornados sean estrictamente los más relevantes y no sea necesario retornar más.

Por otro lado, *Yahoo* y *Bing* tienden a entregar más resultados en este tipo de consultas, porque probablemente sus algoritmos de relevancia sean menos efectivos que *Google* y por lo tanto para suplir esa deficiencia es necesario entregar más documentos para cubrir todo el espectro de sitios relevantes.

En cuanto a *Inquiro*, como es un buscador que está recién construyendo su índice de búsquedas, evidentemente tiende a tener menos documentos que retornar, pero aun así, tampoco devuelve tantos documentos, lo que puede indicar que su algoritmo de relevancia pueda ser más competitivo que los que posee *Yahoo* y *Bing*.

Comparación en base al tiempo de respuesta

Si bien *Yahoo* es el que más respuestas entrega, también es el que más demora en retornar una respuesta, ya que demora 0.9 y 1.31 segundos respectivamente en estas consultas. Aproximadamente 40 ms más que los demás buscadores que presentan tiempos de respuesta similares. Por otro lado, *Inquiro* posee un tiempo de respuesta muy similar a *Google*, pero sigue siendo el que menor tiempo de respuesta tiene de los cuatro.

En la figura (18) también se puede observar que los tiempos de respuesta de los buscadores para ambas consultas son muy similares. La única diferencia es que los tiempos de la consulta referentes al *eclipse solar 2019* son mayores por ser mayor el número de documentos retornados. Lo que es lógico y esperado.

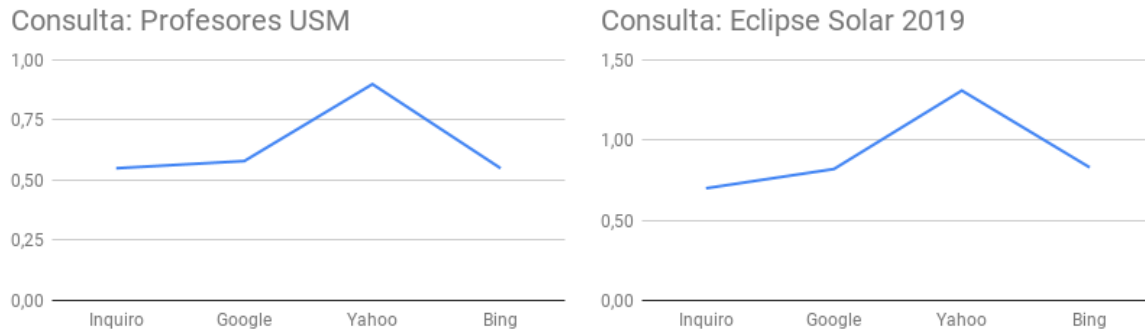


Figura 18: Comparación del tiempo de respuesta de dos consultas informacionales

En conclusión, *Inquiرو* posee el menor tiempo de respuesta en este tipo de consultas, seguido por *Google*, pero claramente como *Inquiرو* posee una colección más reducida, el tiempo de procesamiento también se reduce, lo que influye considerablemente en el desempeño.

4.1.2. Comparación de consultas navegacionales

En esta sección se intenta comparar el comportamiento de los motores de búsqueda al ingresar dos consultas del tipo navegacional: *radio carolina* y *red transantiago*. La primera consulta claramente tiene la intención de llegar al sitio web de la radio carolina y la segunda, llegar al sitio de buses red que operan como transporte en Santiago.

Al igual que la sección anterior, las métricas de comparación utilizadas son el número de resultados que retornan y el tiempo que demoran en hacerlo, además de realizar un análisis cualitativo de resultados.

Comparación en base al número de Resultados

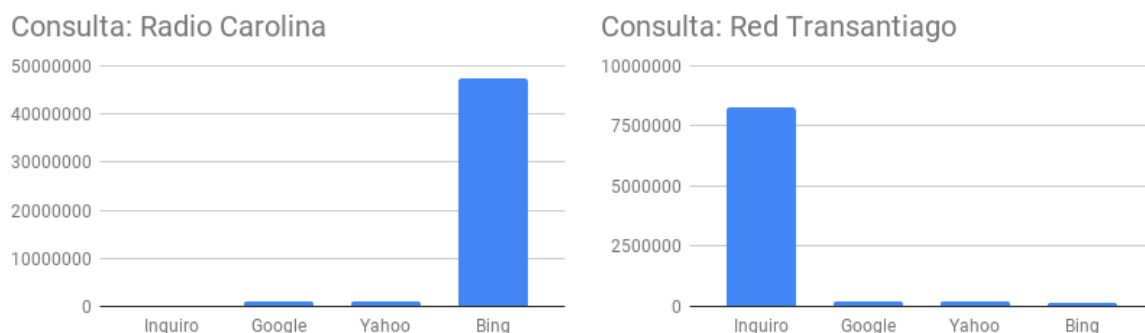


Figura 19: Comparación del N° de resultados de dos consultas navegacionales

En cuanto al número de resultados, se puede visualizar claramente en la figura (19), que el comportamiento de los buscadores en este tipo de consultas navegacionales es muy dife-

rente al de las consultas informacionales mencionadas anteriormente.

En la primera consulta de la *radio carolina*, *Google* y *Yahoo* retornan un número similar de resultados, 922.000 y 992.000 respectivamente. *Inquiro* por su lado, retorna menos de la mitad de esta cantidad, con un valor de 236.232 documentos.

Pero lo más sorpresivo es la diferencia enorme que existe entre *Bing* y los demás, ya que entrega un valor de 47.300.000 de resultados, casi 50 veces más que sus similares.

Esta diferencia también sucede en la segunda consulta, de la *red transantiago*, pero esta vez, es *Inquiro* el que entrega una cantidad muy elevada de documentos.

Se piensa que esto se debe a que los buscadores con los algoritmos presentados en esta memoria no se comportan muy bien con las consultas navegacionales, ya que la intención principal de algoritmos como la *frecuencia de términos*, explicado en el capítulo 2, es la de poder contar la cantidad de palabras que existen en el documento y que fueron ingresadas en la consulta, por lo que se presupone que estos términos se relacionan con el contenido principal del sitio más que del sitio mismo, lo que no es el caso de este tipo de consultas.

Comparación en base al tiempo de respuesta

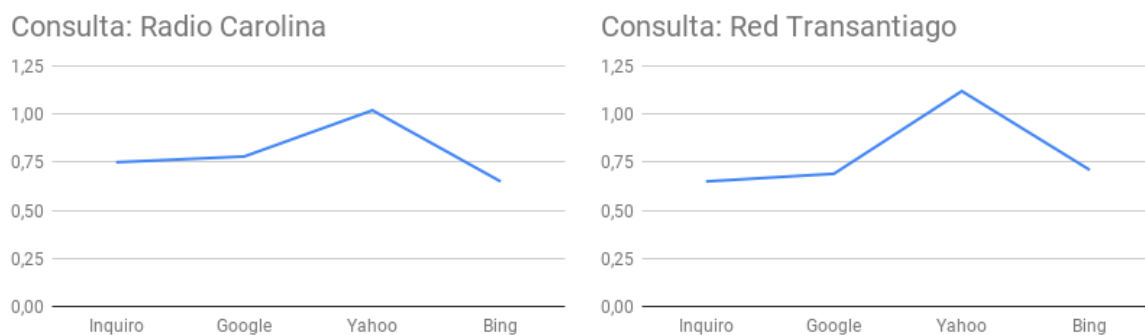


Figura 20: Comparación del tiempo de respuesta de dos consultas navegacionales

El tiempo de respuesta de estas dos consultas también son similares a las anteriores, pero presentan algunas diferencias, que se pueden observar en la figura (20). En la primera consulta, de la *radio carolina*, el buscador *Bing* esta vez posee el menor tiempo de respuesta, pero esto es curioso, ya que también es el que retorna más cantidad de resultados, lo que es contradictorio.

De igual forma, en la segunda consulta, *Inquiro* es el que entrega más cantidad de resultados pero también es el que presenta menos tiempo de respuesta, es decir, el mismo comportamiento anterior.

Se cree que esto se debe a que en ambos casos, los algoritmos utilizados para la recuperación de documentos llegan a un estado en que ya no realizan tantos filtros de búsqueda, si no que retornan un rango mayor de documentos sin filtrar que podrían llegar a ser relevantes para la consulta.

En conclusión, se deduce que este tipo de consulta es de los más difíciles para los buscadores, especialmente si los términos de búsqueda no contienen directamente el nombre del sitio o dominio. Por lo tanto, no hay claridad de qué buscador se comporta mejor en este caso.

4.1.3. Comparación de consultas transaccionales

Por último, en esta sección se intenta comparar el comportamiento de los motores de búsqueda al ingresar dos consultas del tipo transaccional: *comprar caliente camas* y *13c vivo online*. La primera consulta claramente tiene la intención de comprar un producto que en este caso es un caliente camas y en la segunda, acceder al servicio de televisión en vivo del canal nacional 13c²¹.

Al igual que la sección anterior, las métricas de comparación utilizadas son el número de resultados que retornan y el tiempo que demoran en hacerlo, además de realizar un análisis cualitativo de resultados.

Comparación en base al número de Resultados

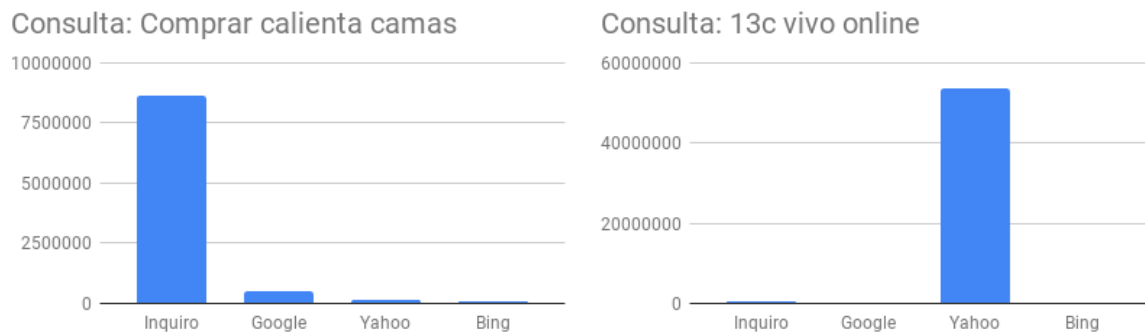


Figura 21: Comparación del N° de resultados de dos consultas transaccionales

Si se compara el número de resultados obtenidos en los cuatro buscadores para estas consultas, se puede visualizar en la figura (21) que el comportamiento en este caso es muy similar al caso de las consultas navegacionales, ya que claramente en ambas consultas existe un buscador que retorna una cantidad mucho mayor que los demás.

En la primera consulta, de *comprar caliente camas* se observa que *Inquirol* retorna una cantidad de resultados de 8.623.309, valor mucho mayor que los demás buscadores que entregan menos de 500.000 documentos.

Lo mismo sucede para la segunda consulta de *13c vivo online*, pero esta vez es *Yahoo* el que devuelve una cantidad de resultados de 53.600.000, valor extremadamente alto en comparación al resto de buscadores.

²¹www.13.cl/c

Al igual que en las consultas anteriores, se piensa que esto se debe a que los buscadores con los algoritmos presentados en esta memoria tampoco se comportan muy bien con las consultas transaccionales, debido a algoritmos como la *frecuencia de términos*, que por definición intentan poder contar la cantidad de palabras que existen en el documento y que fueron ingresadas en la consulta, pero esta técnica presupone que estos términos se relacionan con el contenido principal del sitio más que del servicio que ofrecen, lo que no es el caso de este tipo de consultas.

Comparación en base al tiempo de búsqueda

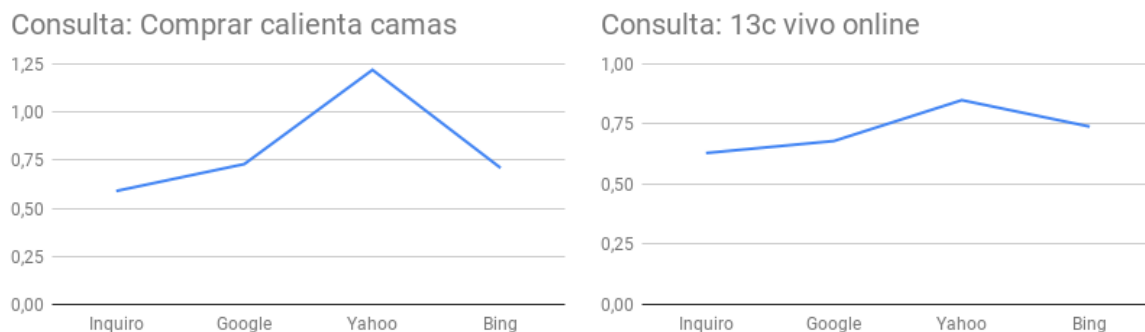


Figura 22: Comparación del tiempo de respuesta de dos consultas transaccionales

En cuanto al tiempo de respuesta de este tipo de consultas transaccionales, el comportamiento también es similar, pero aún así presentan detalles que los diferencian.

En la figura (22) se puede ver que en la primera consulta sucede lo mismo que en la sección anterior, *Inquiro* posee el menor tiempo de respuesta pero a su vez el mayor número de resultados. Además, nuevamente *Yahoo* es el que demora más en dar una respuesta, con un valor de 1.22 segundos, es decir, más del doble que los demás buscadores.

En la segunda consulta los tiempos de respuesta son mucho más similares que en los casos anteriores, pero siempre manteniendo el mismo patrón, *Inquiro* es el buscador que demora menos y *Yahoo* el más lento.

En el primer caso, también se cree que los algoritmos utilizados para la recuperación de documentos llegan a un estado en que ya no realizan tantos filtros de búsqueda, si no que solo retornan un rango mayor de documentos sin filtrar que podrían llegar a ser relevantes para la consulta, lo cual toma menos tiempo de procesamiento.

En el segundo caso, se cree que *Yahoo* se acerca más en los tiempos de respuesta a sus competidores por la misma razón anterior, no se utilizan tantos filtros y se retornan documentos en un rango más amplio de relevancia.

En conclusión, se deduce que este tipo de consulta también es difícil de abordar para los buscadores, especialmente si los términos de búsqueda sólo contienen la acción que quiere

realizar el usuario en el sitio y no un término presente en el sitio mismo. Por lo tanto, no hay claridad de qué buscador se comporta mejor en este caso.

4.1.4. Comparación cualitativa de Resultados

Si se comparan los documentos recuperados de cada buscador y su orden de relevancia, se pueden observar varias diferencias. Esto es útil para analizar el comportamiento de cada motor de búsqueda y saber que es lo que privilegia cada uno a la hora de realizar búsquedas de términos.

A continuación, se analizan cualitativamente los resultados entregados de cada buscador para las distintas consultas mencionadas anteriormente, y se comentan sus diferencias y sus semejanzas. El detalle completo de los resultados entregados por cada buscador puede verse en el anexo A.

Consulta: profesores usm

Al ser esta una consulta del tipo informacional, los primeros resultados entregados por los distintos buscadores, deben hacer referencia a responder a la información requerida. En este caso, la información que se requiere conocer tiene relación con conocer a los profesores de la Universidad Técnica Federico Santa María.

Los primeros resultados de *Inquiro* para esta consulta, hacen relación al sitio oficial de la universidad y luego sólo entrega resultados sobre noticias de la universidad que tienen relación con académicos o profesores.

El buscador *Google*, en cambio, muestra como primeros resultados el sitio directo de la universidad en donde se muestran detalles de los docentes. Luego de esto, muestra el directorio completo de personas y luego empieza a mostrar los docentes separados por cada departamento, como lo es el departamento de mecánica, además de retornar sitios de personas de otras sedes de la misma universidad.

Yahoo por otro lado, muestra como primeros resultados el directorio de personas completo, luego el mismo sitio de docentes USM y también empieza a mostrar resultados de docentes separados por departamento. Pero en este caso, le da prioridad al departamento de eléctrica.

Por último, *Bing* también presenta el directorio completo de personas, pero prioriza más el sitio de la comunidad de la universidad, incluso más que la página de los docentes. Luego al igual que los demás, también presenta los sitios de personas por departamento.

En resumen, *Inquiro* parece priorizar más las noticias relacionadas con la consulta que los demás buscadores. Esto puede deberse a que su algoritmo se basa en la *frecuencia de términos*, es decir, en estas noticias aparecen más veces los términos de la consulta ingresada.

Consulta: eclipse solar 2019

Al ser esta una consulta del tipo informacional, los primeros resultados entregados por los distintos buscadores, deben hacer referencia a responder a la información requerida. En este caso, la información que se requiere conocer tiene relación con el eclipse solar que se produce en este presente año.

Al igual que en la consulta anterior, *Inquiro* solo prioriza sitios de noticias, en primer lugar del canal nacional *24 Horas*²² y luego sitios como *ahoranoticias* y *chile es tuyo*. En todos estos resultados la consulta ingresada aparece tal cual en el título, es decir, aparece en el mismo orden y sin ninguna palabra antes, lo que potencia en gran medida el algoritmo de relevancia de este buscador, pero que es suficiente para retornar documentos con información relevante sobre la consulta.

Google en cambio, presenta como primer resultado un sitio de turismo y luego el sitio nacional oficial del eclipse solar. Después solo empieza a priorizar sitios de noticias también, como el canal *teletrece* y *conicyt*, entre otros. En todos estos es posible encontrar información del eclipse solar.

Por su parte, *Yahoo* también prioriza sitios de turismo, noticias y también de definiciones como *wikipedia*²³. Pero en estas priorizaciones también considera sitios que no son de la web chilena, a pesar de que se haya especificado como región Chile. Esto demuestra que la verticalidad de este buscador no es tan fuerte, ya que los resultados se tienden a escapar del dominio ".cl".

Por último, *Bing* prioriza mucho más las definiciones de *wikipedia* y luego las de turismo y noticias. Si bien este buscador no se sale tanto del dominio chileno, si aparecen en sus resultados sitios ".com", lo que nuevamente indica una tendencia a salirse de la red nacional.

En resumen, todos los buscadores muestran resultados similares, pero algunos como *Yahoo* y *Bing* se tienden a salir del dominio, por ser una consulta que involucra fuertemente dominios internacionales.

Consulta: radio carolina

Al ser esta una consulta del tipo navegacional, los primeros resultados entregados por los distintos buscadores, deben intentar entregar las mejores opciones para llegar a un sitio en particular. En este caso, el sitio al que se quiere llegar es a la página oficial de la radio carolina, una radio nacional FM que es bien popular.

Al consultar el buscador *Inquiro*, este entrega como primer resultado el sitio web de la radio carolina, seguida de subdominios del mismo sitio. Como se encuentra fácilmente el sitio requerido, se puede decir que el buscador se comporta bien para consultas de este tipo.

²²<https://www.24horas.cl/>

²³www.wikipedia.org

Por otro lado, *Google* también presenta como primera opción el sitio requerido, pero las siguientes opciones corresponden a emisoras generales que igualmente se enlazan a la radio carolina, por lo tanto cumple bien su función de navegación.

Yahoo, por su parte también entrega el sitio oficial de la radio carolina como primera opción, luego muestra emisoras que se enlazan a esta radio, pero existe una particularidad, ya que estas emisoras la mayoría no pertenecen al dominio chileno, por lo que nuevamente se estaría perdiendo verticalidad.

Algo muy similar sucede con *Bing*, ya que también entrega el sitio buscado como primera opción, pero luego muestra emisoras chilenas y extranjeras que enlazan con esta radio, además de sitios de definición como wikipedia.

En conclusión, la mayoría de estos buscadores cumplen con el objetivo de llegar al sitio esperado, pero algunos presentan fallas en el dominio que manejan, lo cual afecta la verticalidad requerida.

Consulta: red transantiago

Al ser esta una consulta del tipo navegacional, los primeros resultados entregados por los distintos buscadores, deben intentar entregar las mejores opciones para llegar a un sitio en particular. En este caso, el sitio al que se quiere llegar es a la página oficial del nuevo servicio de transporte red, antes llamado transantiago.

Al analizar los resultados de *Inqui*, se puede observar que el sitio deseado se encuentra en la tercera posición de relevancia y que se presentan dos sitios de noticias primero, anunciando el nuevo servicio de transporte. Si bien este buscador no fue tan efectivo esta vez, el presentar la web deseada entre los tres primeros puestos, es suficiente para decir que cumple la misión navegacional.

Los demás resultados son una mezcla entre sitios de noticias y subdominios del sitio oficial de transporte.

Por otro lado, *Google* entrega el sitio objetivo como primer resultado, luego las páginas antiguas del transantiago y luego sitios de noticias que entregan información con respecto a este medio de transporte. Por lo tanto, cumple con el objetivo de navegación.

Yahoo también muestra como primer resultado el sitio deseado, lo que cumple con el objetivo de navegación, pero si se analizan los demás resultados obtenidos después de este, se puede visualizar que en la segunda opción se muestra un sitio de tecnología que se sale del dominio chileno, y esto se repite más de una vez, lo cual es grave ya que esto sucede incluso explicitando el dominio chileno requerido.

Por último, *Bing* también entrega como primer resultado el sitio deseado, incluso muestra un subdominio del mismo sitio como segunda opción, pero después le siguen sitios de noticias que también se salen del dominio ".cl".

En conclusión, *Google*, *Yahoo* y *Bing* entregan el sitio objetivo como primer resultado, demostrando una clara ventaja en este tipo de consultas, pero los últimos dos se tienden a salir un poco del dominio chileno, lo que les resta efectividad. En este sentido, *Inquiro* tampoco se queda atrás, ya que si bien no entregó el sitio deseado en primer lugar, no se sale del dominio en ningún momento, lo que lo hace un competidor desafiante para *Google*.

Consulta: comprar caliente camas

Al ser esta una consulta del tipo transaccional, los primeros resultados entregados por los distintos buscadores, deben intentar mostrar las mejores opciones para realizar una acción o transacción en la web. En este caso, la acción que se requiere realizar es la de comprar un caliente camas en algún sitio que permite realizar pagos online.

En primer lugar, *Inquiro* entrega como resultados sitios de *Mercadolibre*, un sitio web de compras bastante popular que permite comprar y vender productos libremente a través de la web, y en este caso los resultados corresponden a personas que venden caliente camas en esta página web. También se presentan más sitios de compras como *Buscalibre* y *Sodimac*, además de sitios de noticias que hablan sobre camas o cosas calientes. Así que podría decirse que este buscador sí cumple con el objetivo de compra.

Google por su parte, entrega variados sitios de compras ofreciendo como producto caliente camas. Entre los resultados se encuentran sitios como *Falabella*, *Sodimac*, *Abcdin*, entre otros. Finalizando también con sitios de noticias relacionados con este producto. Este buscador también cumple con la misión de ofrecer opciones de compra.

Por otro lado, *Yahoo* también presenta una gran variedad de tiendas online, como *Hites*, *Abcdin*, *Sodimac*, entre otras. Pero que, al igual que en las consultas pasadas, también entrega resultados fuera del dominio chileno, la mayoría de ellos con sitios informativos, lo que lo hace menos preciso, pero igual de útil.

Y *Bing*, al igual que *Inquiro*, también entrega como primer resultado al sitio *Mercadolibre*, y otras tiendas populares, pero luego de esto nuevamente los resultados se salen del dominio chileno. Lo que no quita el hecho de que cumpla con el objetivo de compra.

En conclusión, todos los buscadores evaluados cumplen con el objetivo de comprar un caliente camas, ya que todos ofrecen variados sitios que permiten realizar el pago de este producto. Teniendo en cuenta nuevamente que *Yahoo* y *Bing* se tienden a salir del dominio chileno en sus búsquedas, a pesar de que se explicita el dominio chileno.

Consulta: 13c vivo online

Al ser esta una consulta del tipo transaccional, los primeros resultados entregados por los distintos buscadores, deben intentar mostrar las mejores opciones para realizar una acción o transacción en la web. En este caso, la acción que se requiere realizar es la de poder ver vía *streaming* un canal nacional de cultura en vivo llamado 13c.

Los resultados de *Inquiro* son bastante buenos, ya que como primer resultado entrega el sitio oficial del canal que ofrece el servicio de *streaming*, pero los demás resultados que ofrece sólo referencian servicios de radio y noticias que hablan del canal de alguna forma. Por lo que se puede asumir que el servicio solicitado se cumple parcialmente.

Por otro lado, *Google* también presenta como primer resultado el mismo sitio oficial del canal solicitado, pero además de esto, presenta sitios adicionales que también ofrecen la posibilidad de ver el mismo canal en vivo, es decir, este buscador ofrece más opciones para poder realizar la acción solicitada.

Yahoo por su parte, entrega casi las mismas páginas de *streaming* mencionadas anteriormente pero en un orden diferente, por ejemplo, en el primer lugar no se encuentra precisamente el sitio oficial del canal, si no que uno bien similar, además de presentar guías online del canal, que nuevamente se salen del dominio chileno, pero que igual logran cumplir con el objetivo transaccional.

Por último, *Bing*, comete el mismo error que *Yahoo*, ya que presenta el sitio oficial del canal equivocado como primera opción, pero que igual presenta el correcto en los siguientes resultados. Además, también presenta más opciones para obtener el servicio de *streaming* pero a la vez saliéndose nuevamente del dominio chileno indicado, lo que claramente le resta precisión.

En conclusión, todos los buscadores ofrecen buenas alternativas para poder ver el canal 13c vía *online*, por lo que todos cumplen con el objetivo. Eso sí algunos presentan más variedad que otros, por ejemplo, *Inquiro* que si bien presenta el sitio correcto como primera opción, no entrega muchas otras alternativas, pero limitándose a estar solamente en el dominio chileno. A diferencia de *Yahoo* y *Bing*, que tienden a salirse bastante.

4.2. Estadísticas de la Colección

El objetivo de esta sección es mostrar y analizar las estadísticas de la colección construida a lo largo de todo el periodo en que el *crawler* diseñado ha estado rastreando información de sitios web activamente.

Estas estadísticas son extraídas de las herramientas que posee el propio motor de búsqueda SolR y además se complementa con la herramienta *Luke*²⁴, que es una herramienta de análisis útil, que tiene la capacidad de acceder a los índices de SolR y mostrar variados análisis de interés.

A continuación, se presentan una serie de estadísticas relevantes para corroborar que la colección construida es representativa de la web chilena y cumple con los diferentes objetivos establecidos para esta memoria.

²⁴<http://www.getopt.org/luke/>

4.2.1. Tiempo de Recolección

En primer lugar, el tiempo de recolección es importante, ya que conocer esto permite realizar estimaciones relevantes de cuanto tardaría el *crawler* actual en rastrear toda la web de Chile.

El inicio del proceso de rastreo comienza el día 3 de Febrero del presente año y termina alrededor del 25 de Junio, es decir, el tiempo de recolección es de cinco meses aproximadamente.

Hay que tener en cuenta que este tiempo de recolección es una estimación aproximada, ya que en las etapas tempranas del desarrollo del *crawler*, este evidentemente era más lento y no presentaba todos los campos que se presentaron en la tabla (5) en el capítulo anterior.

4.2.2. Tamaño de la Colección

El tamaño de la colección construida es esencial para determinar si esta es representativa de la web chilena y además cumple con la función de buscar contenidos relevantes para la población de este país.

En total se han rastreado e indexado 11.656.190 documentos diferentes hasta la fecha del 25 de Junio. Una cantidad superior a la de 10 millones, meta establecida inicialmente en el comienzo de esta memoria. Por lo que se concluye que el proceso de rastreo de sitios ha sido exitoso en ese sentido.

Si bien el tamaño total de la web chilena no se sabe con exactitud, la cantidad recopilada en esta memoria permite deducir que la web es al menos de más de 12 millones de páginas y contando, ya que es evidente que este número aumenta cada día más con el crecimiento de la tecnología.

4.2.3. Cantidad de sitios según tipo de documento

Otra estadística interesante es la cantidad de sitios rastreados según su tipo de documento, es decir, cuántos sitios rastreados corresponden a páginas HTML usuales y cuantos a otro tipo de documento con extensiones como PDF, DOC, PPT y XLS. El resumen de estas cantidades junto al porcentaje de representación en la colección se presentan en la tabla (10).

Se puede apreciar que la inmensa mayoría de los documentos presentes en la colección corresponden al tipo HTML, ya que representa el 99.868 % de toda la colección y luego le siguen los documentos con extensiones PDF, DOC, PPT y XLS, que representan menos del 1 %. Esto deja en evidencia que en general la web chilena no contiene tantos de estos tipos de archivos con extensiones.

| Extensión | Cantidad | Porcentaje |
|--------------|-------------------|--------------|
| HTML | 11.640.758 | 99.867 % |
| PDF | 14339 | 0.123 % |
| DOC | 730 | 0.006 % |
| PPT | 298 | 0.003 % |
| XLS | 65 | 0.001 % |
| Total | 11.656.190 | 100 % |

Tabla 10: Consultas aleatorias diseñadas para la comparación de los buscadores.
Fuente: Elaboración Propia.

Además, dentro de los documentos HTML, es posible saber cuántos de estos poseen protocolos del tipo HTTP y cuantos al tipo HTTPS.

SolR entrega un valor de 1.452.246 documentos del tipo HTTP y 10.203.944 documentos del tipo HTTPS, y esto considerando sólo la red chilena rastreada hasta este momento.

Se sabe que los documentos del tipo HTTPS, se diferencian de los demás por poseer un certificado que puede ser del tipo SSL, que significa que toda la transferencia de datos que se produzca entre los usuarios finales y el servidor *Inquiro* vaya encriptado. De esta forma, si esta transferencia es interceptada de algún modo, de todas formas no podrá ser leída.

Esta diferencia en el número de documentos del tipo HTTPS evidencia claramente la realidad que se vive hoy en día en la web chilena. En un mundo globalizado, en el que la internet es cada vez más usada, la necesidad de proteger los datos aumenta, y por lo tanto también aumenta el uso de certificados en los sitios web.

4.2.4. Profundidad Promedio

La profundidad promedio, como se ha mencionado en los capítulos anteriores, siempre fue pensada para análisis estadístico.

Este valor representa que tan lejos llega el *crawler* en una ejecución de rastreo normal y sin interrupciones. Pero, hay que tener en cuenta que este rastreador ha sido ejecutado varias veces a lo largo del proceso de construcción de la colección, lo que provoca que cada vez que se comienza nuevamente a rastrear, el contador de profundidad se reinicia y empieza a contar nuevamente hasta que se detiene su ejecución.

Según la herramienta SolR, la profundidad promedio de ejecución tiene un valor de 36718 páginas alcanzadas, es decir, en promedio cada url está separada 36718 enlaces de cualquier otra.

También es destacable la mayor profundidad alcanzada entre todas las ejecuciones, llegando a un valor máximo de 104.430 niveles.

4.2.5. Estadísticas con herramienta Luke

Las estadísticas anteriores fueron obtenidas por el propio motor de búsqueda SolR, pero también existen algunos datos interesantes que se pueden obtener con la herramienta *Luke*.

Esta herramienta también es desarrollada por *Apache*, por lo que tiene un buen comportamiento con los índices construidos por *SolR*.

Los datos adicionales más importantes que puede entregar esta herramienta son los siguientes.

Número de Términos de la Colección

Sin duda para un motor de búsqueda, contar con un buen índice de términos es esencial para las búsquedas de consultas, ya que esto es lo que le da el poder de recuperación de documentos relevantes gracias a los algoritmos que se han estudiado en detalle en esta memoria.

Según la herramienta *Luke*, el número total de términos de la colección construida para *Inquiror* hasta la fecha es de 27.442.191 términos. Los cuales corresponden a términos existentes en el diccionario español pero que ya han sido pre-procesadas con técnicas como *Stemming* y *Stopwords Removal*, que se explicaron en detalle en el capítulo 2.

Número de Dominios (Con atributo Domain)

Otra estadística interesante es el número de dominios que existen. Hay que recordar que una url de un sitio posee un dominio, pero un dominio posee muchas url.

Si bien no todos los dominios tienen el mismo número de urls, es posible identificar y cuantificar cuál de estos dominios son más grandes, y que por ende poseen más conexiones. De esta forma, se puede visualizar que sitios tienen más importancia en la red chilena.

Como el campo *domain* fue introducido en una etapa bien tardía del proyecto, solo se contabiliza un total de 3598 dominios diferentes. Los cuales evidentemente no corresponden a el total de dominios de la colección, sino que solo a los que se pueden contabilizar.

Términos más frecuentes según tipo de campo

Al aplicar el pre-procesamiento explicado anteriormente, las palabras resultantes se reducen a palabras como "tiend" o "siti", que muchas veces no poseen las ultimas letras, ya que estas últimas corresponden a las que les dan el género y poseen muchas variaciones en la lengua española. En otras palabras, esto permite llevar estos términos a su palabra raíz y así agrupar todas estas variaciones en un solo término.

A continuación, se presentan los términos más frecuentes según campo indexado, que fueron recopilados con la herramienta *Luke*.

Cuerpo (Body)

En la tabla (11) se presentan los 50 términos más frecuentes en el campo *body* o "cuerpo" de la colección. Es fácil visualizar que el término más popular es "chil", que corresponde al término raíz de todas las derivaciones del país Chile y sus derivaciones como chileno, chilena, chileno, entre otros.

Al ser esta una colección netamente chilena es esperable que este término sea el más frecuente, pero el segundo no es tan esperable, ya que se trata solamente del número "1", que corresponde al segundo término más frecuente.

Los demás términos en la lista corresponden a los más populares en la web chilena en general, muchas de estas pertenecen netamente a términos usuales en internet como "sitios", "cuentas", "facebook", "twitter", entre otros.

La frecuencia de aparición de estos términos varía entre en el rango de 6.941.891 a 10.507.382 veces, solamente contando el texto extraído del cuerpo de cada sitio web.

Título (Title)

En la tabla (12) se presentan los 50 términos más frecuentes en el campo *title* de la colección. También, es fácil visualizar que los términos más populares son "libr", "mercad" y "chil", que aparecen 6.800.000 veces aproximadamente.

El primer término puede corresponder a la raíz de palabras como "libre", "libra", entre otras. El segundo término corresponde a palabras como "mercado", "mercadería" o "mercadillo". Y el tercer término justamente corresponde al más frecuente del campo "Cuerpo", que se mencionó anteriormente.

Los demás términos parecen corresponder a palabras más usadas en los títulos y que generalmente se usan con el fin de captar la atención del lector e invitarlo a seguir leyendo el sitio web.

Este listado de palabras reafirma la premisa de que la mayoría de sitios chilenos corresponden a tiendas comerciales que intentan vender algún producto o servicio. Palabras como "mercad", "vendedor", "arriend" hacen referencia a diferentes agentes de compra y tener este tipo de palabras dentro de los cincuenta términos más frecuentes quiere decir que es una situación que se repite muchas veces en esta red.

Texto Ancla (Anchor Text)

En la tabla (13) se presentan los 50 términos más frecuentes en el campo *anchor text* de la colección. También, es fácil visualizar que los términos más populares son "preci", "menor" y "mayor", que aparecen más de 400.000 veces aproximadamente.

El primer término puede corresponder a la raíz de palabras como "precio", "preciado", en-

| Nº | Frecuencia | Término | Nº | Frecuencia | Término |
|----|------------|-------------|----|------------|-------------|
| 1 | 10507382 | chil | 26 | 8138005 | libr |
| 2 | 9140031 | 1 | 27 | 8072923 | ingres |
| 3 | 9124816 | siti | 28 | 8067081 | punt |
| 4 | 8976951 | condicion | 29 | 7821783 | resultad |
| 5 | 8970257 | compr | 30 | 7820211 | trabaj |
| 6 | 8908180 | 2019 | 31 | 7579015 | descarg |
| 7 | 8883317 | cuent | 32 | 7546110 | menor |
| 8 | 8882359 | termin | 33 | 7540411 | app |
| 9 | 8696795 | ayud | 34 | 7466359 | busqued |
| 10 | 8684073 | informacion | 35 | 7395282 | centr |
| 11 | 8666320 | tiend | 36 | 7360386 | seman |
| 12 | 8560960 | preci | 37 | 7309799 | copyright |
| 13 | 8554667 | social | 38 | 7253722 | 2 |
| 14 | 8540254 | comprar | 39 | 7244788 | seguridad |
| 15 | 8522784 | politic | 40 | 7205510 | gratis |
| 16 | 8516736 | © | 41 | 7180845 | solo |
| 17 | 8507846 | red | 42 | 7153232 | publicación |
| 18 | 8384861 | ofert | 43 | 7140473 | favorit |
| 19 | 8384052 | pago | 44 | 7033169 | crea |
| 20 | 8357839 | ordenar | 45 | 7007189 | youtube |
| 21 | 8332900 | mayor | 46 | 6993910 | historial |
| 22 | 8273559 | facebook | 47 | 6981391 | oficial |
| 23 | 8213897 | twitter | 48 | 6975710 | envi |
| 24 | 8150588 | privacidad | 49 | 6956084 | mercad |
| 25 | 8146461 | categori | 50 | 6941891 | tendenci |

Tabla 11: Los 50 términos más frecuentes del campo *Cuerpo (Body)*
Fuente: Elaboración Propia.

tre otras, que hacen referencia a la cantidad de dinero necesario para adquirir productos o servicios. El segundo y tercer término corresponden a palabras que hacen referencia a un menor o mayor valor de las cosas.

Hay que tener en cuenta que los términos que aparecen en este campo generalmente corresponden a texto que suele describir un enlace a otro sitio. Por ejemplo, los términos "menor" y "mayor" suelen usarse en sitios con paginación de elementos, es decir, en la parte inferior de la página se encuentran enlaces con estos nombres para volver a la página anterior o seguir adelante. O también puede corresponder a sitios en los que se necesita bajar o subir el valor del número de productos, número de personas, entre muchas otras posibilidades.

Otra observación interesante es notar que las palabras usadas en el texto ancla suelen ser más cortas en longitud que las de otros campos. La mayoría de estos términos tiene de uno a cinco letras de largo, incluyendo incluso varios números, lo que es usual en este tipo de

| Nº | Frecuencia | Termino | Nº | Frecuencia | Término |
|----|------------|--------------|----|------------|-----------------|
| 1 | 6870496 | libr | 26 | 461578 | valparais |
| 2 | 6787053 | mercad | 27 | 454561 | aire |
| 3 | 6734336 | chil | 28 | 445508 | 100.000 |
| 4 | 2157508 | cuot | 29 | 442317 | inmuebl |
| 5 | 1702173 | vendedor | 30 | 440051 | fren |
| 6 | 1368942 | nuev | 31 | 437649 | arriend |
| 7 | 1167614 | mejor | 32 | 437419 | calzad |
| 8 | 1160410 | publicad | 33 | 418894 | abs |
| 9 | 1126751 | usad | 34 | 417749 | alarm |
| 10 | 1050113 | envi | 35 | 415737 | acondicionad |
| 11 | 935235 | metropolitan | 36 | 402851 | centr |
| 12 | 933041 | rm | 37 | 390608 | oficin |
| 13 | 614190 | equip | 38 | 387444 | plen |
| 14 | 588464 | catalog | 39 | 381198 | 20.000 |
| 15 | 576063 | 12 | 40 | 371220 | santiag |
| 16 | 574543 | movistar | 41 | 363732 | electric |
| 17 | 570704 | puert | 42 | 361966 | calendar |
| 18 | 568627 | inter | 43 | 361017 | km |
| 19 | 560766 | aut | 44 | 351663 | 15.000 |
| 20 | 537443 | publicacion | 45 | 348795 | 5 |
| 21 | 530958 | camionet | 46 | 344470 | neumatic |
| 22 | 522470 | accesori | 47 | 338037 | beb |
| 23 | 518213 | sicex | 48 | 335705 | acuari |
| 24 | 512253 | gratis | 49 | 329905 | termocalefactor |
| 25 | 463622 | capacitacion | 50 | 326695 | 35.000 |

Tabla 12: Los 50 términos más frecuentes del campo *Título (Title)*

Fuente: Elaboración Propia.

enlaces.

En la tabla (14) se presentan los 50 términos más frecuentes en el campo *domain* de la colección. Este campo corresponde al nombre por el cual se hacen llamar los sitios web en general. Por lo que los términos más frecuentes en este caso, corresponden a los sitios que contienen más subdominios dentro de la colección de *Inquiro*, pero que tampoco corresponden necesariamente a los sitios más grandes, ya que hay que recordar que este campo se agregó en una etapa tardía de la memoria, por lo que el efecto de esta estadística se puede evidenciar de mejor forma a medida que crezca la colección de ahora en adelante.

Se puede visualizar que los términos más populares son "corporacionlosri", "catalogobiblioteca.uach", "teatrochillan" y "hushpuppi", que aparecen más de 140.000 veces aproximadamente. Esto quiere decir que estos sitios son los más grandes actualmente y que a la vez poseen el campo *domain*.

| Nº | Frecuencia | Termino | Nº | Frecuencia | Término |
|----|------------|------------|----|------------|--------------|
| 1 | 910402 | preci | 26 | 46773 | calcular |
| 2 | 525792 | menor | 27 | 46676 | ver |
| 3 | 507600 | mayor | 28 | 46552 | quitar |
| 4 | 415467 | relevant | 29 | 42857 | 10 |
| 5 | 180000 | 2 | 30 | 41028 | lleg |
| 6 | 173074 | 1 | 31 | 40994 | perfil |
| 7 | 147115 | login | 32 | 40618 | z |
| 8 | 135579 | registr | 33 | 40314 | cuand |
| 9 | 132403 | tod | 34 | 40171 | libr |
| 10 | 112604 | vendedor | 35 | 36475 | portabilidad |
| 11 | 110271 | mejor | 36 | 36389 | busqued |
| 12 | 108273 | articul | 37 | 35001 | cost |
| 13 | 107204 | as | 38 | 34747 | hoy |
| 14 | 106838 | veja | 39 | 33312 | deud |
| 15 | 106837 | opção | 40 | 32998 | bloque |
| 16 | 105564 | plan | 41 | 31844 | gris |
| 17 | 103135 | nuev | 42 | 30825 | dat |
| 18 | 93841 | 3 | 43 | 30165 | gb |
| 19 | 93411 | line | 44 | 29814 | publicad |
| 20 | 64090 | 4 | 45 | 29078 | marc |
| 21 | 62032 | eliminar | 46 | 28696 | mpx |
| 22 | 60488 | 5 | 47 | 28540 | tiemp |
| 23 | 56281 | renovacion | 48 | 28387 | http |
| 24 | 55508 | 15 | 49 | 27905 | envi |
| 25 | 53327 | nombr | 50 | 27578 | alto |

Tabla 13: Los 50 términos más frecuentes del campo *Texto Ancla* (Anchor text)
Fuente: Elaboración Propia.

Algunas observaciones interesantes son, que al ser estos nombres de sitios web, tienden a ser más largos que los términos analizados anteriormente.

También si se analiza el rubro de cada dominio se puede concluir que los sitios más grandes en la web chilena corresponden sitios del gobierno, de noticias como páginas de periódicos o radios, o de algún sitio que ofrece productos o servicios.

En conclusión, en este capítulo se presenta como todo el trabajo realizado en la construcción del motor de búsqueda *Inquiro* logra resolver el problema de fondo, que es la no existencia de un buscador centrado en la web chilena.

Se realizaron comparaciones con buscadores actuales que se acercan a resolver el problema, pero que muchas veces se escapan del dominio chileno. *Inquiro* comparado con buscadores como *Google*, *Yahoo* y *Bing* ha demostrado ser igual de efectivo a pesar del poco tiempo de funcionamiento y con una colección más reducida, pero con la particularidad de ser comple-

| Nº | Frecuencia | Termino | Nº | Frecuencia | Término |
|----|------------|-------------------------------|----|------------|-----------------------|
| 1 | 173509 | corporacionlosri | 26 | 7367 | fullbridg |
| 2 | 169010 | catalogobiblioteca.uach | 27 | 6790 | duna |
| 3 | 168404 | teatrochillan | 28 | 6780 | diario.uach |
| 4 | 149438 | hushpuppi | 29 | 6293 | mundomg |
| 5 | 70025 | lcb | 30 | 6287 | triumfo.lanacion |
| 6 | 69037 | melissaoficial | 31 | 6279 | form.elmostrador |
| 7 | 52805 | casaide | 32 | 5886 | autolocal |
| 8 | 48830 | ipad.dafiti | 33 | 5300 | musicapopular |
| 9 | 48670 | dafiti | 34 | 5183 | radio.uchil |
| 10 | 39361 | ahoranotici | 35 | 4833 | festivaldevinachil |
| 11 | 32159 | radioallen | 36 | 4702 | armarket |
| 12 | 32080 | m.dafiti | 37 | 4233 | cursos.universi |
| 13 | 28754 | elmostrador | 38 | 4130 | eldesconciert |
| 14 | 23803 | theclinic | 39 | 4071 | eldinam |
| 15 | 23075 | mega | 40 | 3877 | sodimac |
| 16 | 16727 | matrimoni | 41 | 3869 | zankyou |
| 17 | 15646 | m.elmostrador | 42 | 3835 | plataformaarquitectur |
| 18 | 15642 | tripadvisor | 43 | 3801 | chilev |
| 19 | 15547 | planetadelibr | 44 | 3780 | diarioelcentr |
| 20 | 14677 | leylobby.gob | 45 | 3396 | autologi |
| 21 | 13369 | aqua | 46 | 3094 | prochile.gob |
| 22 | 11932 | gobiernotransparentechile.gob | 47 | 2913 | super45 |
| 23 | 11450 | oan | 48 | 2898 | hoyensantiag |
| 24 | 11379 | lubricantesonline | 49 | 2885 | servel |
| 25 | 10466 | conam | 50 | 2804 | fotech |

Tabla 14: Los 50 términos más frecuentes del campo *Dominio (Domain)*

Fuente: Elaboración Propia.

tamente chileno.

Además, se muestran estadísticas generales de la colección construida con el fin de comprobar los objetivos planteados al principio de esta memoria, como lo es el objetivo de construir una colección representativa de la web chilena, el cual se logra cumplir eficazmente.

CAPÍTULO 5

CONCLUSIONES

En esta memoria, se desarrolla un trabajo muy completo en la aplicación de la informática a problemas reales de la sociedad.

En primer lugar, se realiza una definición del problema central, en el que se explica por qué esto es un problema y merece ser resuelto o mitigado. En este caso, la problemática es la no existencia de un buscador netamente chileno, es decir, que no existe un buscador que solo se concentre en el dominio ".cl".

Para esto es que se construye *Inquiro*, el primer motor de búsqueda que posee una colección completamente chilena. Con esto se logra mitigar en gran parte la necesidad de algunos investigadores, que no están conformes con las búsquedas generales, y que realizan la mayoría de buscadores que son populares actualmente como *Google*, *Yahoo* o *Bing*.

La construcción de este motor de búsqueda, significó el diseño y la implementación de cuatro grandes componentes. El primero es el *crawler*, un programa que se encarga de recorrer la web y que al mismo tiempo, a través de una serie de reglas, va recopilando información de cada sitio rastreado.

El parseador es el segundo componente, que se encarga de extraer el contenido de las secciones específicas de cada documento, como lo son las etiquetas HTML, los títulos, subtítulos, cuerpo del documento, extracción de hyperlinks, entre otros.

Después, viene el *Índice*, que se encarga de almacenar documentos en bruto en alguna estructura de datos eficiente para la posterior recuperación de ese contenido. Más aún, existe un tipo de índice muy especial llamado el *índice invertido* que almacena su colección en una estructura de datos que relaciona cada término en específico con una lista de todos los documentos que contienen dicho término, por lo que recuperar esta lista de documentos para un término en particular tomaría un tiempo lineal constante. Esto es importante porque la eficiencia de este índice es utilizada en la mayoría de buscadores y que actualmente también utiliza *Inquiro*.

Por otro lado está el *Buscador* como tal, que al utilizar un índice invertido trata de devolver de manera eficiente la lista de documentos relevantes para cada término ingresado en las consultas.

Esto se logra a través de una "búsqueda por relevancia" que consta de tres pasos: Realizar una búsqueda de vocabulario para recuperar las listas invertidas de cada término, luego combinar los resultados de estas listas y finalmente ejecutar un puntaje de ranking para cada uno de los documentos de forma que se ordenen del documento más relevante al menos relevante.

El sistema de ranking de *Inquiro* ha demostrado su eficiencia y eficacia en esta memoria y

por qué su algoritmo es de los más utilizados hoy en día por los buscadores populares.

El algoritmo de ranking mencionado corresponde al modelo de vectores espaciales, específicamente a una variación del esquema de Frecuencia de Términos (TF), que representa cada documento como un vector de pesos, en el que cada peso w_{ij} de un término t_i se representa como el número de veces que aparece en el documento d_j .

La variación corresponde al esquema de Frecuencia Inversa de Términos (TF-IDF), que es usado en la mayoría de buscadores, dada su eficiencia y representación. Este esquema se basa principalmente en que cada término en el índice posee su lista de documentos en el que este término aparece, de forma que si se ingresa ese término en una consulta, solo basta con retornar esta lista de documentos asociada.

Para el diseño de la solución de la memoria, se construye una arquitectura que se basa en aprovechar el poder de procesamiento que puede brindar el motor de búsqueda *SolR*, que a su vez permite almacenar un índice invertido eficiente y que además permite la búsqueda de consultas complejas, aplicando técnicas de ranking y combinación de resultados.

También se utilizan técnicas de pre-procesamiento como *Stemming*, que se encargan de llevar cada término almacenado a su palabra raíz, de forma que se puedan agrupar todas las palabras pertenecientes a la misma familia en un solo término. Por otro lado se realiza el procesamiento de *Stopwords Removal*, que remueve ciertas palabras que no aportan significado en las consultas, como lo son los conectores.

Además, una gran parte de la solución consta del desarrollo del *crawler* antes mencionado y que se basa en el lenguaje *Python*. Que junto a la librería *Scrapy* se construye un rastreador recursivo capaz de recorrer toda la web chilena, extrayendo su información y actualizando su índice al mismo tiempo.

Es por esto que se decide utilizar el *framework Django*, ya que este está basado en *python* y permite diseñar páginas web robustas rápidamente.

Para el diseño del *crawler* como tal, se desarrollan dos scripts principales, *crawler.py* y *main.cpp*, el primero en lenguaje *Python* y el segundo en *C++*.

crawler.py contiene todos los métodos y funcionalidades de un rastreador, es decir, tiene la capacidad de realizar llamadas http a los sitios web de Chile, parsear su información para extraer de manera focalizada el texto deseado, y de almacenar cada documento junto a su índice invertido en la colección de *SolR*.

Además, este componente tiene la capacidad de actualizar el índice ya construido, de forma que se puedan re-visitarse los sitios rastreados previamente, para actualizar los términos asociados a ese documento. Esto está pensado para sitios que se consideren demasiado antiguos y que la información que poseen ya no concuerda con la rastreada en un principio.

En el caso del parseo de la información, se pueden clasificar en dos grandes tipos: El parseo

de documentos HTML y el parseo de archivos con extensión.

El parseo de documentos HTML es el más comúnmente usado en los rastreadores, y consiste en parsear las etiquetas que posee la estructura HTML. Esto se logra gracias a una librería llamada *Tika*, que está escrita en *Java* pero con una API en *Python*, que posee la capacidad de extraer todo el texto de documentos, soportando una multitud de extensiones.

Las etiquetas más importantes para rastrear en este tipo de documentos son el título (*title*), el cuerpo (*body*), el texto ancla (*anchor text*), la url y el dominio de cada uno de los sitios rastreados.

También en el caso del parseo de archivos con extensión, primero se verifica que extensión posee ese archivo, ya sea PDF, XLS, DOC o PPT, y luego se extrae todo el texto detectado en cada uno de ellos con la ayuda de la librería *Tika*.

Finalmente, el texto extraído a través de los dos tipos de parseo, es procesado y almacenado en el índice invertido de *SolR*.

Por otro lado se tiene el componente *main.cpp*, que contiene todos los comandos del sistema en general. A través de un *Makefile* permite ejecutar diferentes acciones como reiniciar la base de datos de *SolR*, importar archivos de información para el índice invertido, ejecutar el *crawler*, ejecutar el actualizador de índices, realizar chequeos de archivos, hacer limpieza de archivos y también combinaciones de estas acciones.

En otras palabras, *main.cpp* actúa como un orquestador que gestiona todas las funcionalidades del sistema construido para el motor de búsqueda *Inqui*ro.

Entre los componentes más destacados de *main.cpp* se encuentran el *Controlador*, *SolR*, *Crawler.py* y *Spdlog*.

El Controlador se encarga de controlar las principales acciones que se pueden realizar. Además de realizar acciones como la verificación de carpetas, aplicar sistemas de *Logging* como *Spdlog*, entre otros. Es por esto que dependiendo de la acción a realizar, se comunica con los sub-componentes correspondientes.

Por ejemplo, si el controlador recibe un comando de acción que solicita reiniciar la colección o importar archivos de información al índice, el Controlador debe comunicarse con *SolR* para ejecutar estas acciones.

En el caso contrario, como lo es el rastreo de sitios o generaciones de grafo de la red chilena, el controlador debe comunicarse con *crawler.py*, sub-componente explicado anteriormente y que es controlado desde *main.cpp*. Su principal función es rastrear y extraer información de sitios web y con esto generar archivos con extensión CSV que posteriormente pueden ser indexados a la colección final.

Todos estos componentes contribuyen a la construcción de la colección final del buscador *Inqui*ro, pero este también debe cubrir otros aspectos, como por ejemplo el diseño de su interfaz y que tan fácil de usar es esta plataforma.

Es por esto que se diseña una interfaz pensada en los chilenos, en el público objetivo de este buscador. Con un diseño de colores que es fácil de identificar, los cuales son rojo, azul y blanco, los colores patrios de Chile. Además de esto, se siguen reglas de usabilidad al mostrar resultados de consultas, ya que colores como azul, verde y gris facilitan la búsqueda de información.

En cuanto a la implementación de la solución, se utilizan recursos de *software* como *Sublime Text 3*, *SolR*, *Django*, *Gitkraken* y *Ubuntu*. Por otro lado, se utilizan dos recursos de *hardware*, un notebook y un servidor de alta capacidad provisto por la Universidad Técnica Federico Santa María.

Los alcances y limitaciones de esta memoria están dados por el avance de la web chilena y la tecnología respectivamente, ya que sin importar la dimensión de la red de Chile, *Inquiro* puede continuar indexando páginas indefinidamente. En cuanto a las limitaciones, se puede mencionar el avance de la tecnología, ya que se intentan usar las mejores herramientas que se encuentran disponibles hasta la fecha para cada uno de los componentes de este motor de búsqueda, pero es bastante probable que en un tiempo más estas tecnologías queden obsoletas y la eficiencia actual ya no sea suficiente.

Con respecto a los resultados entregados por esta memoria se pueden mencionar varios. Uno de los resultados principales es que la creación de un motor de búsqueda netamente chileno es posible y es más, sirve como una herramienta poderosa para la población chilena que tiene interés por el contenido nacional y que los buscadores actuales no son suficientes para satisfacer sus necesidades de información.

Otro gran resultado son las estadísticas obtenidas gracias a la construcción de la web chilena, ya que actualmente no se tenía un número aproximado de la dimensión de esta red, pero que ahora puede estimarse de mejor forma. De esta misma manera las estadísticas nos entregan información valiosa de lo que se encuentra en la web chilena, se puede apreciar claramente que existe una predominancia de sitios comerciales y bancarios en esta red, que las palabras más frecuentes corresponden a términos que se relacionan con compras, tiendas y precios.

También se pueden concluir con las estadísticas, que el número de documentos con extensiones como PDF, DOC, PPT y XLS es muy bajo comparado con el número de documentos HTML indexados, es decir, hay predominancia de sitios web planos que no contienen estos tipos de documentos y que solo entregan información a través de páginas estáticas.

Al evaluar el comportamiento del buscador *Inquiro* con respecto a otros buscadores como *Google*, *Yahoo* y *Bing* se concluye que las búsquedas del tipo informacional son las más adecuadas para *Inquiro*, ya que el esquema de la frecuencia de términos (TF) utilizado como algoritmo de *ranking* para la relevancia de documentos según una consulta, se comporta mejor con este tipo de búsquedas que intentan conseguir información en base a términos.

Los objetivos de esta memoria eran claros y desafiantes, consistía en la implementación de principio a fin de un motor de búsqueda que sea capaz de navegar por la red chilena y que

permita buscar contenidos de forma fácil y con exactitud. Es un objetivo más que válido, ya que no se encuentra implementado actualmente en Chile y contribuye a suplir la necesidad de conocimiento presente en el país.

Las principales contribuciones y aplicaciones del trabajo realizado para esta memoria son muchas. La implementación del primer buscador vertical exclusivamente chileno marca un antes y un después en el ámbito de los buscadores nacionales. Se puede aplicar a situaciones en las que se necesite una búsqueda nacional focalizada en contenidos como documentos con extensión o en situaciones en que se requieran resultados no convencionales para ciertas consultas.

Este trabajo además, es un gran aporte para Chile y para la investigación en el área de la recuperación de la información (IR), ya que provee más que una plataforma de búsqueda, sino que también abre una nueva puerta para la investigación de la red chilena y su comportamiento.

Por esto mismo, el trabajo realizado debe motivar a los interesados en las búsquedas web. Para ellos se recomienda nunca dejar de probar nuevos métodos o técnicas que permitan obtener nuevas métricas y nuevas conclusiones para contribuir con la investigación, ya que cada nuevo conocimiento impacta directamente en la toma de decisiones de los chilenos y en una mejora en sus vidas.

Para todas aquellas personas que se sientan motivadas por la investigación en el área de la recuperación de la información, se les recomienda que diseñen ustedes mismos los rastreadores, índices y motores de búsqueda, ya que esto permite aprender realmente lo que es un sistema tan complejo como los buscadores. Hoy en día, en que la sociedad cada vez es más dependiente de la tecnología, estos motores de búsqueda juegan un papel crucial, ya que la velocidad del flujo de la información es cada vez más rápido y esto se debe a los sistemas que permiten encontrar sitios web más rápidamente, como *Inquiro*.

Finalmente, hay que recalcar que el trabajo realizado en esta memoria ha sido muy completo, llegando incluso a compararse con sistemas tan sofisticados como lo son los buscadores que llevan un buen tiempo en el mercado. Esto también abre un marco de trabajo para futuras investigaciones de la red chilena y los documentos que posee.

Perfectamente se puede ampliar el trabajo que se hizo con *Inquiro*, para estudiar o mejorar cada uno de sus componentes. Por ejemplo, una de las ampliaciones propuestas sería la investigación de nuevos algoritmos de *ranking* de documentos, que permitan una mejor evaluación en cuanto a las necesidades de los chilenos. Se puede mejorar bastante lo que es la combinación de resultados al evaluar relevancia, por lo que se recomienda que las futuras investigaciones apunten en esta dirección.

Quizás se puede mejorar la arquitectura, la eficiencia de resultados con el uso de cache o quizás la creación de perfiles de usuario que realizan búsquedas permitan mejorar la calidad de los resultados, ya que estos poseen búsquedas en común entre sí.

ANEXOS

ANEXO A: RESULTADOS DE CONSULTAS BUSCADORES

6.1. Resultados Inquiro

6.1.1. Consulta: profesores usm

Universidad Técnica Federico Santa María

<https://usm.cl/>

¿ Admisión Investigación Vida Universitaria Comunidad USM Universidad Campus y Sedes Diplomas Directorio Trabaja en USM Contacto English ver...

DSC_0048 • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/05/17/usm-realiza-muestra-informativa-en-vina-del-m...>

¿ Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

DSC_0043 • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/05/17/usm-realiza-muestra-informativa-en-vina-del-m...>

¿ Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

DSC_0040 • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/05/17/usm-realiza-muestra-informativa-en-vina-del-m...>

¿ Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

USM reconoce a profesores con excelencia en docencia de pregrado en Viña del Mar • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/12/26/usm-reconoce-a-profesores-con-excelencia-en-d...>

¿ Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Curso de vida saludable realiza trekking recreativo para profesores • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2019/06/05/curso-de-vida-saludable-realiza-trekking-recr...>

¿ Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Carrera T. U. en Mecánica Industrial de la USM realiza charla sobre equipos de procesa-

miento en Minería • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/05/22/carrera-t-u-en-mecanica-industrial-de-la-usm-...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Departamento de Mecánica de la USM realiza innovador seminario en mantenimiento industrial • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/12/27/departamento-de-mecanica-de-la-usm-realiza-in...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Profesora Caren Carmona inicia perfeccionamiento docente • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/05/04/profesora-caren-carmona-inicia-perfeccionamie...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Comité paritario de la USM Viña del Mar participa de capacitación en la UPLA • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/12/27/comite-paritario-de-la-usm-vina-del-mar-parti...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Departamento de Electrotecnia e Informática de la USM realiza charla “Introducción a Raspberry Pi” • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/12/13/departamento-de-electrotecnia-e-informatica-d...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

USM realiza masivo operativo integral de salud para sus funcionarios en Viña del Mar • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/09/10/usm-realiza-masivo-operativo-integral-de-salu...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Sansanos visitan instalaciones de Chilquinta Energía en Quilpué • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/06/25/sansanos-visitando-instalaciones-de-chilquinta-...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

USM realiza exitoso seminario sobre Manufactura Digital Avanzada en Viña del Mar • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/06/26/usm-realiza-exitoso-seminario-sobre-manufactu...>

🔍 Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

nidades Responsabilidad Social Campus y Sedes Diplom...

Profesor Cristian Ahumada recibe título de Doctor en Ciencias de la Ingeniería mención Automática • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/06/29/profesor-cristian-ahumada-recibe-titulo-de-do...>

Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Carrera de Técnico Universitario en Proyecto y Diseño Mecánico de la USM renueva su acreditación por seis años • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2019/01/18/74519/>

Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Director de Vinculación con el Medio de la USM en Viña del Mar invita a participar de enriquecedora propuesta metodológica a profesores • USM Noticias • Universidad Técnica Federico Santa

María <https://noticias.usm.cl/2018/05/03/director-de-vinculacion-con-el-medio-de-la-us...>

Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

USM se adjudica proyecto que permitirá actualizar la carrera de TU en Mecánica Automotriz • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/07/05/usm-se-adjudica-proyecto-que-permitira-actual...>

Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

Con degustación de comida saludable culmina taller de Súper-Alimentos en la USM en Viña del Mar • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2018/07/11/con-degustacion-de-comida-saludable-culmina-t...>

Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

USM desarrollará programas académicos dirigidos a profundizar en la prevención de riesgos e higiene industrial • USM Noticias • Universidad Técnica Federico Santa María

<https://noticias.usm.cl/2019/05/02/usm-desarrollara-programas-academicos-dirigid...>

Noticias Arte y Cultura Ciencia y Tecnología Deporte Docencia Economía y Negocios Humanidades Responsabilidad Social Campus y Sedes Diplom...

6.1.2. Consulta: eclipse solar 2019

Eclipse Solar 2019 - 24horas

<https://www.24horas.cl/eclipsesolar2019/>

Toggle navigation 24HORAS.CL Eclipse Solar 2019 Señales en vivo Síguenos Nacional Política

Primarias Internacional Estados Unidos BBC Mundo ...

ECLIPSE SOLAR 2019: ¿POR QUÉ SE ORIGINA EL FENÓMENO ASTRONÓMICO? | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/tendencias/260439-eclipse-solar-2019-en-ch...>

Ahora es Noticia TNE Sistema frontal Tornado Los Angeles Selección Chilena Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de...

Eclipse solar 2019 en Chile - Astronomía en Chile y Astroturismo

<https://www.chileestuyo.cl/eclipse-solar-2019-en-chile-lugares-para-el-astroturi...>

Menu ¿Qué hacer? Astronomía Bienestar y relax Cultura y patrimonio Deporte y aventura Naturaleza Turismo de negocios Vida urbana Vinos y com...

ECLIPSE SOLAR 2019: CONSEJOS PARA VER DE FORMA SEGURA EL FENÓMENO ASTRONÓMICO | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/tendencias/260276-eclipse-solar-2019-en-ch...>

Ahora es Noticia TNE Sistema frontal Tornado Los Angeles Selección Chilena Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de...

ECLIPSE SOLAR 2019: CONOCE LOS LUGARES DONDE PODRÁS VER EL FENÓMENO ASTRONÓMICO EN O'HIGGINS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/nacional/260197-eclipse-solar-2019-en-chil...>

Ahora es Noticia TNE Sistema frontal Tornado Los Angeles Selección Chilena Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de...

ECLIPSE SOLAR 2019: CONOCE LOS MEJORES LUGARES PARA VER EL EVENTO EN LA REGIÓN DE VALPARAÍSO | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/nacional/259976-eclipse-solar-2019-en-chil...>

Ahora es Noticia TNE Sistema frontal Tornado Los Angeles Selección Chilena Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de...

ECLIPSE SOLAR 2019: CONOCE LOS MEJORES LUGARES PARA VER EL FENÓMENO QUE OCURRIRÁ EN JULIO | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/nacional/259177-eclipse-solar-en-chile-con...>

Ahora es Noticia Mineros Atrapados Cambio de Gabinete Lluvia en Santiago Javiera Suárez Noticias Programas Dato Útil Nacional Mundo Tendenci...

ECLIPSE SOLAR 2019: CONOCE CUÁNTO DURARÁ Y DÓNDE VER EL FENÓMENO ASTRONÓMICO EN ATACAMA | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/nacional/259208-eclipse-solar-2019-en-chil...>

Ahora es Noticia Compin Mineros Atrapados Copa América 2019 Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida Deportes ...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=solar>

Ahora es Noticia TNE Sistema frontal Tornado Los Angeles Selección Chilena Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=ps&page=3>

Ahora es Noticia Puerto Montt Cyberday 2019 Instituto Nacional Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida Deport...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=Eclipse%20solar>

Ahora es Noticia Donald Trump Miguel Fuente-Alba Venezuela Caburgua Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida D...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=eclipse%20solar>

Ahora es Noticia Donald Trump Miguel Fuente-Alba Venezuela Caburgua Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida D...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=solar&page=2>

Ahora es Noticia TNE Sistema frontal Tornado Los Angeles Selección Chilena Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=Regi%C3%B3n%20de%20Coquimbo>

Ahora es Noticia Puerto Montt Cyberday 2019 Instituto Nacional Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida Deport...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=IPS>

Ahora es Noticia Donald Trump Miguel Fuente-Alba Venezuela Caburgua Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida D...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=ps&page=2>

Ahora es Noticia Puerto Montt Cyberday 2019 Instituto Nacional Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida Deport...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=Regi%C3%B3n%20de%20Coquimbo&page=2>

Ahora es Noticia Puerto Montt Cyberday 2019 Instituto Nacional Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida Deport...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=regi%C3%B3n%20de%20Atacama>

Ahora es Noticia Compin Mineros Atrapados Copa América 2019 Noticias Programas Dato

Útil Nacional Mundo Tendencias Calidad de vida Deportes ...

AHORA NOTICIAS | Ahoranoticias.cl

<https://www.ahoranoticias.cl/buscador/?tag=regi%C3%B3n%20de%20Atacama&page=1>
Ahora es Noticia Compin Mineros Atrapados Copa América 2019 Noticias Programas Dato
Útil Nacional Mundo Tendencias Calidad de vida Deportes ...

Comité de Asignación de Tiempo chileno en el Telescopio APEX - SOCHIAS

<https://sochias.cl/comite-de-asignacion-de-tiempo-chileno-en-el-telescopio-apex/...>
Skip to content SOCHIAS Sitio oficial de la Sociedad Chilena de Astronomía Nosotros Directiva
Lista de socios Comités Asignación de Tiempo A...

6.1.3. Consulta: radio carolina

Radio Carolina

<http://www.carolina.cl/>
Programas Carolina TV Noticias Música Videos Nerd Buster Concursos Noticias Estos son los
desvíos de tránsito durante la "Maratón de Santiag...

Comunicado Oficial sobre cambios en Radio Carolina : Radio Carolina

<http://www.carolina.cl/noticias/2018/04/comunicado-oficial-sobre-cambios-en-radi...>
Programas Carolina TV Noticias Música Videos Nerd Buster Concursos Noticias Comunicado
Oficial sobre cambios en Radio Carolina 2 de abril de...

radio : Radio Carolina

<http://www.carolina.cl/etiqueta/radio/>
Programas Carolina TV Noticias Música Videos Nerd Buster Concursos radio Noticias 2 de
abril de 2018 Comunicado Oficial sobre cambios en Rad...

Frecuencias : Radio Carolina

<http://www.carolina.cl/frecuencias/>
Programas Carolina TV Noticias Música Videos Nerd Buster Concursos Frecuencias ¡Encuen-
tra la señal Fm de tu ciudad! Santiago - 99.3 Arica - ...

Francisco : Radio Carolina

<http://www.carolina.cl/etiqueta/francisco/>
Programas Carolina TV Noticias Música Videos Nerd Buster Concursos Francisco Lo más visto
en Carolina + Compartidos + Vistos 1 05.04.2019 No...

pega : Radio Carolina

<http://www.carolina.cl/etiqueta/pega/>
Programas Carolina TV Noticias Música Videos Nerd Buster Concursos pega Noticias 9 de
mayo de 2018 Dato de pega: Aerolínea busca dos amigos ...

app : Radio Carolina

<http://www.carolina.cl/etiqueta/app/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos app Noticias 6 de marzo de 2019 Ya no podrán agregarte a grupos de Whatsa...

fotos : Radio Carolina

<http://www.carolina.cl/etiqueta/fotos/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos fotos Cine 26 de marzo de 2019 [FOTOS] Avengers: Endgame confirman a otro...

aplicaciones : Radio Carolina

<http://www.carolina.cl/etiqueta/aplicaciones/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos aplicaciones Noticias 22 de junio de 2017 [VIDEO] Estos son los nuevos em...

Whatsapp : Radio Carolina

<http://www.carolina.cl/etiqueta/whatsapp/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos Whatsapp Noticias 10 de diciembre de 2018 Ahora podrás llamar a los grupo...

fotos : Radio Carolina

<http://www.carolina.cl/etiqueta/fotos/page/29/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos fotos Música 27 de febrero de 2015 Romeo Santos en imágenes Hizo bailar y...

fotos : Radio Carolina

<http://www.carolina.cl/etiqueta/fotos/page/3/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos fotos Noticias 13 de noviembre de 2018 [FOTOS] Así quedó la casa de Miley...

Instagram : Radio Carolina

<http://www.carolina.cl/etiqueta/instagram/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos Instagram Música 27 de marzo de 2019 Paren todo: Justin Bieber se retira ...

app : Radio Carolina

<http://www.carolina.cl/etiqueta/app/page/10/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos app Tecnología 6 de marzo de 2015 ¿Conoces Happn? La app que amenaza el r...

cine : Radio Carolina

<http://www.carolina.cl/etiqueta/cine-2/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos cine Cine 3 de abril de 2019 [VIDEO] Thanos ganó: mira el nuevo spot de A...

video : Radio Carolina

<http://www.carolina.cl/etiqueta/video/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos video Noticias 5 de abril de 2019 [FOTOS] Conoce el nuevo personaje de Pe...

series : Radio Carolina

<http://www.carolina.cl/etiqueta/series/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos series Noticias 1 de abril de 2019 Game Of Thrones tendrá documental Se a...

pelicula : Radio Carolina

<http://www.carolina.cl/etiqueta/pelicula/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos pelicula Cine 3 de abril de 2019 [VIDEO] Thanos ganó: mira el nuevo spot ...

amor : Radio Carolina

<http://www.carolina.cl/etiqueta/amor/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos amor Música 5 de marzo de 2019 Cami Gallardo rompió el silencio: "Yatra e...

cine : Radio Carolina

<http://www.carolina.cl/etiqueta/cine-2/page/2/>

Programas Carolina TV Noticias Música Videos Nerd Buster Concursos cine Cine 12 de marzo de 2019 [VIDEO] Mira el nuevo trailer de Aladdín La...

6.1.4. Consulta: red transantiago

Red - The Clinic

<https://www.theclinic.cl/etiqueta/red/>

Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home Poder Chilean Nius Webeo Columnas Mundo al instante D...

EN QUÉ CONSISTIRÁ RED": EL NUEVO SISTEMA DE TRANSPORTE CAPITALINO QUE REEMPLAZARÁ AL TRANSANTIAGO | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/nacional/253643-red-cambio-transantiago-mo...>

Ahora es Noticia Donald Trump Miguel Fuente-Alba Venezuela Caburgua Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida D...

Red Metropolitana de Movilidad

<http://www.red.cl/>

Reclamos y sugerencias | Llámanos 600 730 0073 800 73 0073 | Síguenos en: Toggle navigation Inicio Planifica tu viaje Cuando llega Paradas y...

RED, Nuevo Transporte para Santiago. - Fundación Jaime Guzmán

<https://www.fjguzman.cl/2019/04/10/red-nuevo-transporte-para-santiago/>

Fundación Jaime Guzmán Menu Nosotros Quiénes somos Equipo FJG Directorio Área de Servicio Público Área de Formación Área Legislativa Área Co...

Posts con la etiqueta Red" | Duna 89.7

<https://www.duna.cl/tag/red/>

Programas Entrevistas Podcast Noticias Música Concursos EN VIVO / / DUNA TV PODCAST Sintonía Crónica 7 Junio Terapia Chilensis 7 Junio Nada ...

Red": El nuevo nombre del Transantiago que costó una .ªsesoría creativa de alto impacto"de \$550 millones | El Desconcierto

<https://www.eldesconcierto.cl/2019/03/12/red-el-nuevo-nombre-del-transantiago-qu...>

Facebook El Desconcierto Twitter El Desconcierto YouTube El Desconcierto Instagram El Desconcierto Menú El Desconcierto País Política Opinión...

Transantiago | El Dínamo

<https://www.eldinamo.cl/temas/transantiago/>

Cerrar publicidad Cerrar publicidad El Dínamo Nacional Actualidad Educación Opinión Deportes Ambiente Videos Reportajes Mujer Dínamo Faceboo...

Accesibilidad | Red Metropolitana de Movilidad

<http://www.red.cl/acerca-de-red/accesibilidad>

Reclamos y sugerencias | Llámanos 600 730 0073 800 73 0073 | Síguenos en: Toggle navigation Inicio Planifica tu viaje Cuándo llega Paradas y...

Malla Nocturna | Red Metropolitana de Movilidad

<http://www.red.cl/mapas-y-recorridos/malla-nocturna>

Reclamos y sugerencias | Llámanos 600 730 0073 800 73 0073 | Síguenos en: Toggle navigation Inicio Planifica tu viaje Cuándo llega Paradas y...

Dos buses del Sistema Red fueron quemados durante marcha no autorizada | El Dínamo

<https://www.eldinamo.cl/nacional/2019/05/31/dos-buses-del-sistema-red-fueron-que...>

Cerrar publicidad Cerrar publicidad El Dínamo Nacional Actualidad Educación Opinión Deportes Ambiente Videos Reportajes Mujer Dínamo Faceboo...

Indignante: Operadora del Transantiago debe trabajar con su hija porque no tiene quién la cuide - El Mostrador

<https://www.elmostrador.cl/noticias/multimedia/2018/08/07/indignante-operadora-d...>

El Mostrador, el primer diario digital de Chile. Toggle navigation Ingresar BIENVENIDOS Si ya está suscrito, ingrese aquí: Email Contraseña ...

Transantiago | El Desconcierto

<https://www.eldesconcierto.cl/tag/transantiago/>

Facebook El Desconcierto Twitter El Desconcierto YouTube El Desconcierto Instagram El Desconcierto Menú El Desconcierto País Política Opinión...

PRIMER RECORRIDO DE TRANSANTIAGO AL AEROPUERTO SE INAUGURARÁ DOS DÍAS AN-

TES DE LA LLEGADA DEL PAPA A CHILE | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/nacional/212869-primer-recorrido-de-transa...>

Ahora es Noticia Donald Trump Miguel Fuente-Alba Venezuela Caburgua Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida D...

transporte público - The Clinic

<https://www.theclinic.cl/etiqueta/transporte-publico/>

Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home Poder Chilean Nius Webeo Columnas Mundo al instante D...

PARADEROS DIGITALES Y APP CON TIEMPOS DE BUSES: LOS CAMBIOS QUE TRAE RED | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/nacional/253688-paraderos-digitales-app-ti...>

Ahora es Noticia Donald Trump Miguel Fuente-Alba Venezuela Caburgua Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida D...

Trabajadores de Subus de Transantiago realizan un paro de advertencia esta mañana - Duna 89.7 |

Duna 89.7 <https://www.duna.cl/noticias/2016/08/01/trabajadores-de-subus-de-transantiago-re...>

Programas Entrevistas Podcast Noticias Música Concursos EN VIVO / / DUNA TV PODCAST Sintonía Crónica 7 Junio Terapia Chilensis 7 Junio Nada ...

Local Dos Niveles La Dehesa - UF 284 en Mercado Libre

https://local.mercadolibre.cl/MLC-478522112-local-dos-niveles-la-dehesa-_JM

Mercado Libre Chile - Donde comprar y vender de todo Crea tu cuenta Ingresa Tus compras Tu historial Categorías Tiendas oficiales Ofertas de...

Choferes de Transantiago anuncian paro en partido Chile-Holanda - Duna 89.7 | Duna 89.7

<https://www.duna.cl/noticias/2014/06/22/choferes-de-transantiago-anuncian-paro-e...>

Programas Entrevistas Podcast Noticias Música Concursos EN VIVO / / DUNA TV PODCAST Sintonía Crónica 7 Junio Terapia Chilensis 7 Junio Nada ...

Posts con la etiqueta "transporte" | Duna 89.7

<https://www.duna.cl/tag/transporte/>

Programas Entrevistas Podcast Noticias Música Concursos EN VIVO / / DUNA TV PODCAST Información Privilegiada 17 Junio La Tercera PM 17 Junio...

Premios para incentivar a quienes cargan y pagan con tarjeta Bip ¿Por qué no? - Duna 89.7 | Duna

89.7 <https://www.duna.cl/podcasts/premios-para-incentivar-a-quienes-cargan-y-pagan-co...>

Programas Entrevistas Podcast Noticias Música Concursos EN VIVO / / DUNA TV PODCAST Sintonía Crónica 7 Junio Terapia Chilensis 7 Junio Nada ...

6.1.5. Consulta: comprar caliente camas

Caliente Camas 12.000 - \$ 12.000 en Mercado Libre

https://articulo.mercadolibre.cl/MLC-441026311-caliente-camas-12000-_JM

Mercado Libre Chile - Donde comprar y vender de todo Bienvenido Ingresar a tu cuenta para ver tus compras, favoritos, etc. Ingresar Crea tu cu...

Caliente Camas Norwood 1 Plaza \$10000 - \$ 10.000 en Mercado Libre

<https://articulo.mercadolibre.cl/MLC-489427742-caliente-camas-norwood-1-plaza-10...>

Mercado Libre Chile - Donde comprar y vender de todo Bienvenido Ingresar a tu cuenta para ver tus compras, favoritos, etc. Ingresar Crea tu cu...

Cam Jansen Y El Misterio De Los Huesos De Dinosaurio, David A. Adler comprar el libro

<https://www.buscalibre.cl/libro-cam-jansen-y-el-misterio-de-los-huesos-de-dinosa...>

50 % Descuento en Cientos de Libros de Editorial Planeta. VER AQUÍ Inicio Libros Compra en USA Repuestos de autos Venta Empresas Libros Envío...

Camas Usadas - Camas, Usado en Mercado Libre Chile

<https://hogar.mercadolibre.cl/dormitorio/camas/camas-usadas>

Mercado Libre Chile - Donde comprar y vender de todo Crea tu cuenta Ingresar Tus compras Tu historial Categorías Tiendas oficiales Ofertas de...

Camas Usadas - Camas, Usado en Mercado Libre Chile

https://hogar.mercadolibre.cl/dormitorio/camas/usados/camas-usadas_Desde_451

Mercado Libre Chile - Donde comprar y vender de todo Crea tu cuenta Ingresar Tus compras Tu historial Categorías Tiendas oficiales Ofertas de...

Camas Usadas - Camas, Usado en Mercado Libre Chile

https://hogar.mercadolibre.cl/dormitorio/camas/usados/camas-usadas_Desde_401

Mercado Libre Chile - Donde comprar y vender de todo Crea tu cuenta Ingresar Tus compras Tu historial Categorías Tiendas oficiales Ofertas de...

Camas Usadas - Camas, Usado en Mercado Libre Chile

https://hogar.mercadolibre.cl/dormitorio/camas/usados/camas-usadas_Desde_351

Mercado Libre Chile - Donde comprar y vender de todo Crea tu cuenta Ingresar Tus compras Tu historial Categorías Tiendas oficiales Ofertas de...

Camas | Homy.cl

<https://www.sodimac.cl/sodimac-homy/category/cat2940020/Camas>

Bienvenido a Homy.cl. Todo para amoblar y decorar tus espacios. Homy, diseño para todos. Todos los Muebles y la mejor Decoración en 10 cuota...

Camas | Homy.cl

https://www.sodimac.cl/sodimac-homy/category/cat2940020/camas?utm_source=bloghom...

Bienvenido a Homy.cl. Todo para amoblar y decorar tus espacios. Homy, diseño para todos.

Todos los Muebles y la mejor Decoración en 10 cuota...

Camas | El Desconcierto

<https://www.eldesconcierto.cl/tag/camas/>

Facebook El Desconcierto Twitter El Desconcierto YouTube El Desconcierto Instagram El Desconcierto Menú El Desconcierto País Política Opinió...

caliente caliente eo - The Clinic

<https://www.theclinic.cl/etiqueta/caliente-caliente-eo/>

Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home Poder Chilean Nius Webeo Columnas Mundo al instante D...

"Soy Tapsin: caliente día, caliente noche"

<https://www.theclinic.cl/2011/10/18/soy-tapsin-caliente-dia-caliente-noche/>

Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home Poder Chilean Nius Webeo Columnas Mundo al instante D...

Camas y Box Springs Camas Funcionales en Mercado Libre Chile

<https://hogar.mercadolibre.cl/dormitorio-camas-box-springs-funcionales/>

Mercado Libre Chile - Donde comprar y vender de todo Bienvenido Ingres a tu cuenta para ver tus compras, favoritos, etc. Ingres Crea tu cu...

Camas y Box Springs Camas Funcionales para Adultos en Mercado Libre Chile

<https://hogar.mercadolibre.cl/dormitorio-camas-box-springs-funcionales/adultos/>

Mercado Libre Chile - Donde comprar y vender de todo Bienvenido Ingres a tu cuenta para ver tus compras, favoritos, etc. Ingres Crea tu cu...

La Raza Cómica | CAM

<http://razacomica.cl/sitio/tag/cam/>

Search for: A desalambrar Agarrando pueblo Civilización y Barbarie La ciudad letrada Persiana americana Las entrañas del monstruo Piel negra...

Camas colonizadas | El Desconcierto

<https://www.eldesconcierto.cl/2018/10/19/camas-colonizadas/>

Facebook El Desconcierto Twitter El Desconcierto YouTube El Desconcierto Instagram El Desconcierto Menú El Desconcierto País Política Opinió...

camas elásticas - ChileInflable

https://chileinflable.cl/etiqueta-producto/camas-elasticas/?per_row=3&shop_view=...

Facebook Email Instagram YouTube WhatsApp Facebook Email Instagram YouTube WhatsApp 0 Favoritos Somos Inicia tu negocio Productos Fabricació...

camas elásticas - ChileInflable

https://chileinflable.cl/etiqueta-producto/camas-elasticas/?per_page=36

Facebook Email Instagram YouTube WhatsApp Facebook Email Instagram YouTube WhatsApp 0 Favoritos Somos Inicia tu negocio Productos Fabricació...

camas elásticas - ChileInflable

https://chileinflable.cl/etiqueta-producto/camas-elasticas/?per_row=2&shop_view=...
Facebook Email Instagram YouTube WhatsApp Facebook Email Instagram YouTube
WhatsApp 0 Favoritos Somos Inicia tu negocio Productos Fabricació...

camas elásticas - ChileInflable

https://chileinflable.cl/etiqueta-producto/camas-elasticas/?per_row=4&shop_view=...
Facebook Email Instagram YouTube WhatsApp Facebook Email Instagram YouTube
WhatsApp 0 Favoritos Somos Inicia tu negocio Productos Fabricació...

6.1.6. Consulta: 13c vivo online

13C | 13.cl

<http://www.13.cl/c>
T13 Deportes 13 Espectáculos 13C RecTV 13 Radios Oasis FM Play FM Sonar FM Tele13 Radio
Programas Programación Blog El Entrevistatorio Cultu...

CHILE VS. BRASIL EN VIVO: MIRA AQUÍ EL PARTIDO ONLINE | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/deportes/207350-chile-vs-brasil-en-vivo-mi...>
Ahora es Noticia TNE Sistema frontal Tornado Los Angeles Selección Chilena Noticias Progra-
mas Dato Útil Nacional Mundo Tendencias Calidad de...

El fallecido Gustavo Sánchez cuenta su historia en el programa "Vidas" de Sebastián Eyzaguirre

en el 13C <https://www.theclinic.cl/2012/10/31/el-fallecido-gustavo-sanchez-cuenta-su-histo...>
Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home
Poder Chilean Nius Webeo Columnas Mundo al instante D...

Francisco Pérez-Bannen conducirá programa cultural en Canal 13C — Pudahuel

<https://www.pudahuel.cl/noticias/2018/04/francisco-perez-bannen-conducira-progra...>
En Pudahuel escuchas Canciones Bajo las Estrellas Radio Online facebook twitter instagram
youtube facebook twitter instagram youtube MENÚ No...

[EN VIVO] URUGUAY VS. RUSIA: SEÑAL ONLINE, VIDEOS Y MINUTO A MINUTO | Ahoranoticias.cl

<https://www.ahoranoticias.cl/fifaworldcup/noticias/228363-uruguay-vs-rusia-en-vi...>
Publicidad Ahora es Noticia Lluvia en Arica Desborde río Loa Venezuela Noticias Programas
Dato Útil Nacional Mundo Tendencias Deportes 5 Día...

[EN VIVO] BÉLGICA VS. JAPÓN: SEÑAL ONLINE, VIDEOS Y MINUTO A MINUTO | Ahoranoticias.cl

<https://www.ahoranoticias.cl/fifaworldcup/noticias/229111-en-vivo-belgica-vs-jap...>
Publicidad Ahora es Noticia TNE Tornado Los Angeles Selección Chilena Sistema frontal No-

ticias Programas Dato Útil Nacional Mundo Tendencias...

Paula - Online

<http://radio.paula.cl/player/>

HOME Reportajes y Entrevistas Reportajes Entrevistas Belleza Bienestar Cara Cuerpo Pelo Moda Tiempo Libre Artes Visuales Deco Libros Música ...

Radio en Vivo. - Chilean British Radio

http://radio.britanico.cl/?page_id=294

Home Programación Misión y Visión Quiénes Somos Equipo Señal Online Podcast Radio en Vivo. WordPress Theme | Square by Hash Themes

Sigue en vivo el eclipse

<https://www.theclinic.cl/2013/05/09/sigue-en-vivo-el-eclipse/>

Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home Poder Chilean Nius Webeo Columnas Mundo al instante D...

GAME OF THRONES: ASÍ PUEDES VER EN VIVO Y ONLINE EL CAPÍTULO FINAL DE LA SERIE | Ahoranoticias.cl

<https://www.ahoranoticias.cl/noticias/tendencias/261473-final-game-of-thrones-en...>

Ahora es Noticia Compin Mineros Atrapados Copa América 2019 Noticias Programas Dato Útil Nacional Mundo Tendencias Calidad de vida Deportes ...

Señal Online - La Red

<https://lared.cl/senal-online>

Programas Hola Chile Así Somos Aventura 4x4 Intrusos Cara a Cara Mundomotor Cada Día Mejor Expediente S WWE RAW Mentiras Verdaderas WWE SMAC...

Casino online Chile | Mejores casinos - Promociones online - Premios

<https://www.iide.cl>

iide.cl Casino online Chile: mejores casinos y juegos con premio millonario Bienvenido a nuestro site en el que te vamos a acompañar durante...

En vivo Canal 13 | 13.cl

<https://www.13.cl/en-vivo>

T13 Deportes 13 AR13 13C RecTV 13 Radios Oasis FM Play FM Sonar FM Tele13 Radio Programas Capítulos Programación En vivo 13Now En vivo En vi...

INIA » Exitoso proyecto de domesticación de abejorro nativo tendrá espacio en Canal 13C

<http://www.inia.cl/blog/2017/05/29/exitoso-proyecto-de-domesticacion-de-abejorro...>

INTRANET INIA MINAGRI INDAP SAG CONAF CNR ODEPA FIA INIA INFOR CIREN FUCOA ACHIPIA AGROSEGUROS Inicio Nosotros Quiénes Somos Organigrama y E...

Supermercado Online | Lider.cl

<https://www.lider.cl/supermercado/home.jsp>

¿Permitir el uso de cookies para mejorar la experiencia de navegación? Sí No Locales Compras

Anteriores Seguimiento de compras Pickup Ingres...

Slayer, en vivo en Suecia 2011

<https://www.theclinic.cl/2013/04/10/slayer-en-vivo-en-suecia-2011/>

Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home Poder Chilean Nius Webeo Columnas Mundo al instante D...

Vivo Positivo - The Clinic

<https://www.theclinic.cl/etiqueta/vivo-positivo/>

Sobre TC-online Contacto TC en Facebook TC en Twitter TC en Flickr RSS Bares Menú Home Poder Chilean Nius Webeo Columnas Mundo al instante D...

Radio Bío-Bío | Señal Online

<https://vivo.biobiochile.cl/player/?radio=valparaiso>

SIGUE EN VIVO RADIO RALLY MUNDIAL DE BIO BIO DETENIDO Escuchando 0 Presiona Play ELIGE TU SEÑAL Santiago 99.7 Valparaíso 94.5 Concepción 98....

Radio Bío-Bío | Señal Online

<https://vivo.biobiochile.cl/player/?radio=concepcion>

SIGUE EN VIVO RADIO RALLY MUNDIAL DE BIO BIO DETENIDO Escuchando 0 Presiona Play ELIGE TU SEÑAL Santiago 99.7 Valparaíso 94.5 Concepción 98....

Radio Bío-Bío | Señal Online

<https://vivo.biobiochile.cl/player/?radio=santiago>

SIGUE EN VIVO RADIO RALLY MUNDIAL DE BIO BIO DETENIDO Escuchando 0 Presiona Play ELIGE TU SEÑAL Santiago 99.7 Valparaíso 94.5 Concepción 98....

6.2. Resultados Google

6.2.1. Consulta: profesores usm

Unidades académicas y docentes - Universidad Técnica Federico ...

<https://www.usm.cl/universidad/unidades-academicas-docentes/>

USM Transparente Universidad Acreditada 6 Años G9 Universidades Públicas No Estatales. Mecesup Reuna Universia Estudia en Valparaíso Agrupación de ...

Directorio de Personas USM

www.directorio.usm.cl/ escudo UTFSM. Directorio de Personas USM El Directorio en línea de personas y departamentos de la UTFSM ... Profesores y Funcionarios ...

Profesores - mecanica.usm.cl - Universidad Técnica Federico Santa ...

<https://www.mecanica.usm.cl/> Quiénes Somos

Danilo Estay Barrientos. Ingeniero Civil en Metalurgia. Universidad de Atacama. Oficina: C-

333. E-mail: danilo.estay@usm.cl · ver más. fernando labbe ...

Académicos

www.ociv.usm.cl/es/quienes-somos/profesores/

OCIV UTFSM ... Cuerpo Docente. ARANDA N., Vivian. Instructor Académico. ARAYA A., Felipe. Instructor Académico. BAZÁEZ G., Ramiro. Profesor Auxiliar ...

Comunidad USM · Viña del Mar · Universidad Técnica Federico Santa ...

www.vinadelmar.usm.cl/comunidad-usm/

Comunidad USM. Alumnos; Profesores; Funcionarios; Ex-Alumnos; Dirección General de Investigación y Postgrado; Interés general; Redes sociales ...

Profesores - mecanica.usm.cl - Universidad Técnica Federico Santa ...

<https://www.mecanica.usm.cl> › Quiénes Somos

César Morales Velasco. Ph. D., Engineering Mechanics, Virginia Polytechnic Institute & State University, 1997. M. Sc., Engineering Mechanics, Virginia ...

Docentes · Defider · Universidad Técnica Federico Santa María

www.defider.usm.cl/nosotros-central/docentes-central/

ANDRADE VERGARA BYRON. Jefe Área de Extensión/ Educación Física; Fono: 32 2654044; Email: byron.andrade@usm.cl ...

Comunidad USM · San Joaquín · Universidad Técnica Federico Santa ...

www.sanjoaquin.usm.cl/comunidad-usm/

Comunidad USM. Alumnos; Profesores; Funcionarios; Ex-Alumnos; Empresas; Dirección General de Comunicaciones; Interés general; Redes sociales ...

USM reconoce a profesores con excelencia en docencia de pregrado ...

<https://noticias.usm.cl/.../usm-reconoce-a-profesores-con-excelencia-en-docencia-de-p...>

26 dic. 2018 - 26 · diciembre · 2018. El premio reconoce a aquellos profesores que obtuvieron las mejores calificaciones en la evaluación docente ...

Comunidad USM · Vitacura · Universidad Técnica Federico Santa María

www.vitacura.usm.cl/comunidad-usm/

Comunidad USM. Alumnos; Profesores; Funcionarios; Ex-Alumnos; Empresas; Dirección General de Comunicaciones; Interés general; Redes sociales ...

USM reconoce a profesores por su excelencia en docencia de ...

<https://noticias.usm.cl/.../usm-reconoce-a-profesores-por-su-excelencia-en-docencia-d...>

2 ene. 2019 - 2 · enero · 2019. La ceremonia premió a un total de 16 profesores, entre los cuales 13 recibieron la distinción de Maestros Destacados y tres ...

USM reconoce la excelencia de la labor docente de sus profesores ...

<https://noticias.usm.cl/.../usm-reconoce-la-excelencia-de-la-labor-docente-de-sus-profe...>

19 feb. 2019 - Con el objetivo de reconocer la labor de los profesores de todos los Campus y Sedes de la Universidad Técnica Federico Santa María, ...

Profesores • Departamento de Arquitectura USM • Universidad ...

arquitectura.usm.cl/profesoresutfsm/

21 · Agosto · 2017. Profesores de nuestro Departamento se encuentran actualmente trabajando en el desarrollo de dos proyectos FONDART adjudicados en la ...

mecanica.usm.cl - Universidad Técnica Federico Santa María

<https://www.mecanica.usm.cl/>

En el Departamento de Ingeniería Mecánica USM preparamos a nuestros estudiantes ... Nuestros profesores y estudiantes de postgrado llevan a cabo además ...

Profesores Jornada Completa

www.metalurgia.usm.cl/index.php/profesores/profesores-jornada-completa

Área de especialización: Procesos Acuáticos Aplicados a Metalurgia Extractiva y Protección Ambiental. Oficina: C-331 (Casa Central). E-mail: juan.ibanez@usm.

Académicos • Departamento de Electrónica ... - Electrónica USM

www.electronica.usm.cl/personas/nuestro-equipo/academicos/

A continuación se señalan los profesores del departamento. Seleccione un profesor para conocer su información personal.

Profesores - mecanica.usm.cl - Universidad Técnica Federico Santa ...

www.mecanica.usm.cl/10-profesores?start=10

Publicaciones Recientes. Mario Toledo, José Jiménez, Luis Cárdenas, Romain Gers, Jaime Espinoza, Combustion of biofuels-diesel blends in an isothermal ...

Profesores de USM Casa Central recibieron premios por Excelencia ...

<https://www.mecanica.usm.cl> › Ver más noticias

En la tradicional jornada se reconocieron a 12 profesionales con la distinción Maestros de Excelencia y a otros 49 en la categoría Maestros Destacados...

USM rinde homenaje al destacado profesor Walter Grote - USM Noticias

<https://noticias.usm.cl/2019/.../usm-rinde-homenaje-al-destacado-profesor-walter-grot...>

28 may. 2019 - En una emotiva ceremonia en la que participaron profesores, estudiantes y funcionarios, la Universidad Técnica Federico Santa María rindió ...

USM fortalece su planta académica con 37 nuevos profesores • USM ...

<https://noticias.usm.cl/.../usm-fortalece-su-planta-academica-con-37-nuevos-profesore...>

25 mar. 2019 - 25 · marzo · 2019. Fueron recibidos oficialmente en la Jornada de Inducción Institucional de Profesores, instancia en donde se abordaron ...

6.2.2. Consulta: eclipse solar 2019

Eclipse Solar Total 2 de Julio de 2019 - Turismo Astronómico Chile

<https://www.turismoastronomico.cl/eclipse-solar-total-2-de-julio-de-2019.html>

El 2 de julio de 2019 se producirá un eclipse solar total que será visible por una delgada franja del océano Pacífico y parte de América del Sur. La totalidad del ...

Mapa del Eclipse - AURA - Chilean Eclipse 2019

<https://www.eclipsesolar2019.cl/eclipse/mapa-del-eclipse/>

Mapa del Eclipse - 02 de julio de 2019 ... Credito: Observatorio Nacional Solar de EEUU, AURA y la Fundación Nacional de Ciencias de EE.UU. Usando la ...

Eclipse Solar 2019: ¿cuánto cuestan y dónde comprar los lentes ... - T13

<https://www.t13.cl/.../eclipse-solar-2019-cuanto-cuestan-y-donde-comprar-lentes-filtro...>

hace 36 minutos - Se acerca el eclipse solar del 2 de julio y cada día aumenta más el interés por adquirir los lentes certificados para poder observar el fenómeno ...

¿A qué hora se verá el eclipse solar en tu ciudad? | Tele 13

<https://www.t13.cl/.../eclipse-solar-2019-donde-se-vera-ciudades-2-julio-horarios-270...>

hace 17 horas - Revisa el horario del eclipse solar en tu ciudad. El evento será visible de ... Especial Eclipse Total 2019: ingresa a este enlace. El fenómeno se ...

Eclipse solar 2019 - Coquimbo | Eclipses Chile 2019 - 2020 v0 - Conicyt

<https://www.conicyt.cl/eclipseschile1/eclipse-solar-2019-coquimbo/>

Eclipse 2019 - Coquimbo. El martes 2 de julio de 2019, un eclipse total de Sol cruzará la Región de Coquimbo y parte de la Región de Atacama. El eclipse ...

Eclipse Total Solar Archivos - Sernatur Región de Coquimbo - Chile

turismoregiondecoquimbo.cl/que-hacer/eclipse-total/

Región de Coquimbo, el mejor escenario mundial del Eclipse Solar 2019 ... Un Eclipse Solar Total, fenómeno cósmico que será estudiado por los científicos del ...

lo que tienes que saber sobre el evento del siglo: el eclipse solar total ...

<https://www.gochile.cl › Home › Artículos>

Esto es lo que necesitas saber sobre el eclipse solar 2019 Valle del Elqui.

Eclipse Total de Sol 2019 - Astroturismo Región de Coquimbo

<https://astroturismochile.travel/index.php/eclipse-total-de-sol-2019/>

El 2 de julio de 2019 se producirá uno de los eventos astronómicos más espectaculares en el planeta, por lo que se espera que cientos de miles de personas ...

Eclipse Solar Atacama 2019

eclipsesolaratacama19.cl/

Existen eclipses del Sol y de la Luna, que ocurren solamente cuando el Sol y la Luna se alinean con la Tierra de una manera determinada. Esto sucede ...

Eclipse Solar 2019 - 24horas

<https://www.24horas.cl/eclipsesolar2019/>

Sigue estas recomendaciones para visualizar el eclipse cuidando la salud ocular. Los lentes de sol normales sólo filtran los rayos UV e infrarrojos de manera ...

ECLIPSE SOLAR 2019: CUÁNDO ES Y EN QUÉ PARTE DE CHILE ...

<https://www.ahoranoticias.cl/.../263588-eclipse-solar-2019-en-chile-donde-verlo-2-jul...>

5 jun. 2019 - Un eclipse total de Sol ocurre cuando la Luna cubre completamente la cara visible del Sol, proyectando una sombra en la superficie de la ...

Eclipse Solar 2019 - AHORA NOTICIAS | Ahoranoticias.cl

[https://www.ahoranoticias.cl/buscador/?tag=Eclipse %20Solar %202019](https://www.ahoranoticias.cl/buscador/?tag=Eclipse%20Solar%202019)

Eclipse solar 2019: Estos son los mejores lugares para ver el fenómeno en ... Eclipse solar 2019: Esto cuestan los peajes para viajar desde Santiago a la región ...

ECLIPSE SOLAR 2019 / OBSERVATORIO CRUZ DEL SUR - Ticketpro

<https://www.ticketpro.cl/evento/402/eclipse-solar-2019-observatorio-cruz-del-sur>

2 jul. 2019 - OBSERVATORIO CRUZ DEL SUR. ECLIPSE SOLAR 2019 desde 14:00 hasta 19:00 hrs. [https://www.observatoriocruzdelsur.cl ...](https://www.observatoriocruzdelsur.cl...)

Eclipse Solar Total 2019 - OPCC

opcc.cl/eclipse-solar-total-2019/

A poco menos de un año del Eclipse Total de Sol que se producirá en gran parte del área de la Región de Coquimbo el 2 de julio de 2019, los equipos de EPO ...

Eclipse Solar 2019 - Información importante para visitas | ESO Chile

<https://www.eso.org/public/chile/events/astro-evt/solareclipse2019/important-info/>

La visualización del eclipse solar total dependerá del clima, por lo cual ESO no garantiza que el eclipse será visible. ESO organizará actividades alternativas en ...

Lo que hay que saber para no perderse el Eclipse Solar 2019 ...

<https://www.revistaenfoque.cl/lo-que-hay-que-saber-para-no-perderse-el-eclipse-solar-...>

El 2 de julio, un eclipse solar se podrá ver desde casi todo el país. Las regiones de Atacama y Coquimbo tendrán un 100 % de cobertura. Santiago, un 92 %.

eclipse solar 2019 | Diario El Día

www.diarioeldia.cl/tags/eclipse-solar-2019

5 jun. 2019 - Desde la Gobernación del Elqui, junto a los municipios y La Armada, se determinó mantener la prohibición de pernoctar en el borde costero ...

EXPERIENCIA ECLIPSE SOLAR TOTAL - 2 de julio 2019 ...

<https://www.ecoturismolaserena.cl/es.../experiencia-eclipse-solar-total-2-de-julio-2019/>

Eclipse Solar Total es un evento natural único en el mundo a desarrollarse en la Región de Coquimbo el próximo 2 de julio 2019. Le recomendamos que ...

Eclipse Solar 2019 - Elquiglobal

www.elquiglobal.cl/tag/eclipse-solar-2019/

Paihuano prepara diversos servicios y actividades para recibir eclipse solar · ... Vicuña será un lugar privilegiado para observar el eclipse total de sol en 2019.

6.2.3. Consulta: radio carolina

Radio Carolina

www.carolina.cl/

Contacto. Dirección: Avenida Vicuña Mackenna #1370. Ñuñoa, Santiago Chile. Teléfono: (+56 2) 2810 80 10. Email : contacto@carolina.cl ...

Radio Carolina Online | Emisora.cl

<https://emisora.cl/radio-carolina/>

Calificación: 4.8 - 2.052 votos Escuchá Carolina 99.3 FM a través de emisora.cl. Con un simple click puedes escuchar todas las mejores emisoras de radio de Chile.

Radio Carolina Online En Vivo | Emisora.cl

<https://emisora.cl/carolina-santiago/>

Calificación: 4.9 - 390 votos Escuchá Radio Carolina 99.3 FM Santiago a través de emisora.cl. Con un simple click puedes escuchar todas las mejores emisoras de radio de Chile.

Radio Carolina online - Señal en vivo - 99.3 MHz FM, Santiago de ...

<https://onlineradiobox.com> › Chile › Metropolitana › Santiago de Chile

Calificación: 4.5 - 2 opiniones Radio Carolina 99.3 MHz FM, Santiago de Chile, Chile - escuchar radio online gratis en OnlineRadioBox.com.

Radio Carolina - Señal Online

<https://www.eltelon.com/radios/carolina>

Disfruta Radio Carolina online las 24hrs del día. Descubre cientos de radios en vivo en nuestro sitio.

Escuchar Radio Carolina en vivo - Radios de Chile

www.radios-chilenas.com › Chile › Santiago

Calificación: 4.5 - 350 votos Escucha Radio Carolina online e más de 400 radios chilenas. La mejor música, programas y las últimas noticias. Escucha radios de Chile gratis!

Carolina TV En Vivo | TV Online Chile

tv-online.cl/carolina-tv-en-vivo.php

Carolina TV En Vivo, TV Online Chile: Sintoniza todas las senales abiertas de Television ... ricardo campinas sp, brasil...saludos a todos d ela radio carolina.

Mega adquiere radio Carolina para sumar a su conglomerado de ...

<https://www.emol.com/.../Mega-adquiere-radio-Carolina-para-sumar-a-su-conglomerado...>

29 oct. 2016 - El grupo Bethia posee además la radio Candela, el sitio Mega.cl y el ... Los motivos de la elección de radio Carolina fueron comunicados por ...

Movidas en medios de Álvaro Saieh: Copesa vendería Radio Carolina ...

<https://www.eldesconcierto.cl/.../movidas-en-medios-de-alvaro-saieh-copesa-venderia-...>

7 oct. 2016 - Zero y Carolina son las dos radios que el Grupo Dial, perteneciente a Copesa,

vendería al grupo que controla Mega y Radio Canela. Además ...

Radio Carolina confirma que Javier Olivares no sigue y congela.^a ...

<https://www.biobiochile.cl/.../radio-carolina-confirma-que-javier-olivares-no-sigue-y-...>

2 abr. 2018 - Tras el polémico video de aparente connotación sexual protagonizado por Karol Lucero y que comenzó a viralizarse la semana pasada, Radio ...

Radio Carolina lanza una bebida energética - Publimark.cl

<https://www.publimark.cl/marketing/.../radio-carolina-lanza-una-bebida-energetica.ht...>

20 nov. 2018 - Radio Carolina, que busca hace tiempo ser reconocida como la “más prendida”, decidió crear una bebida energética como una manera de ...

Radio Carolina llega a Concepción | La Tercera

<https://www.latercera.com/noticia/radio-carolina-llega-a-concepcion/>

1 feb. 2010 - Radio Carolina crece. Una de las estaciones juveniles más importantes y escuchadas del país amplía su público y hoy, desde las 8 horas, ...

Mega firmó acuerdo para comprar radio Carolina | soychile.cl

<https://www.soychile.cl/.../Mega-firmo-acuerdo-para-comprar-radio-Carolina.aspx>

28 oct. 2016 - Ayer El mandatario usó su cuenta de Twitter para referirse al fenómeno ocurrido en la ciudad de la Región del Biobío. El intendente del Biobío ...

El Tipógrafo » Javier Olivares deja Radio Carolina tras 11 años ...

<https://eltipografo.cl/.../javier-olivares-deja-radio-carolina-tras-11-anos-audidores-relac...>

2 abr. 2018 - Javier Olivares deja Radio Carolina tras 11 años: Auditores relacionaron su salida a las críticas que le lanzó a Karol Lucero. Publicado el 2 ...

Alex Anwandter se lanza en picada contra Radio Carolina y el Pelao ...

<https://www.theclinic.cl/.../alex-anwandter-se-lanza-en-picada-contr-radio-carolina-y-...>

8 mar. 2017 - El cantante Alex Anwandter acusó a Radio Carolina de utilizar una de sus canciones justo en el momento en que tiraban tallas homofóbicas al ...

Javier Olivares dejará Radio Carolina tras 11 años al aire - Cooperativa

<https://www.cooperativa.cl/noticias/.../radio/javier...radio-carolina.../123937.html>

2 abr. 2018 - Después de 11 años al aire, Javier Olivares se despidió de Radio Carolina y puso fin a su programa “Superatómico”. “Hoy cierra la cortina ...

6.2.4. Consulta: red transantiago

Red Metropolitana de Movilidad

<https://www.red.cl/>

Red Metropolitana de Movilidad. ¿Qué es Red? Infórmate aquí sobre la Red Metropolitana de Movilidad · Mapas. Conoce los mapas de los recorridos.

Conoce los Recorridos | Red Metropolitana de Movilidad - DTPM

www.transantiago.cl/mapas-y-recorridos/conoce-los-recorridos

Inicio; Conoce los Recorridos. Conoce los Recorridos. Puedes ver el detalle de tus recorridos. Buscador de recorridos. Código del recorrido. Ejemplo: 402.

Transantiago

[www.transantiago.cl/planifica?start=Grecia %2C+Penalolen&end=Diagonal...](http://www.transantiago.cl/planifica?start=Grecia%2C+Penalolen&end=Diagonal...)

Esta página no puede cargar Google Maps correctamente. ¿Eres el propietario de este sitio web? Aceptar. Cuando Llegas Cuándo Llegas; Paradas y Bip; Cómo ...

Estos son los 17 nuevos recorridos del Transantiago que beneficiarán ...

<https://www.eldesconcierto.cl/.../estos-son-los-17-nuevos-recorridos-del-transantiago-...>

hace 3 días - Este domingo salieron a las calles los nuevos recorridos y trayectos del Transantiago (renombrado como Red por el gobierno del presidente ...

Este domingo debutaron nuevos recorridos del Transantiago ...

<https://www.cooperativa.cl/noticias/pais/.../transantiago/...transantiago/.../065944.html>

hace 3 días - En ese contexto, hay que revisar todos los servicios que presta Red Metropolitana de Movilidad (conocido simplemente como Red)", destacó ...

La ruta que seguirá Red para reemplazar al Transantiago

<https://www.latercera.com/nacional/noticia/la-ruta-seguira-red...transantiago/567502/>

12 mar. 2019 - El Presidente Piñera detalló los alcances de plan, que incluirá nuevas flotas de buses eléctricos a nivel nacional, la licitación para renovar las ...

De Transantiago a Red: ¿Solución real o mera estrategia ...

<https://radio.uchile.cl/.../de-transantiago-a-red-solucion-real-o-mera-estrategia-comuni...>

12 mar. 2019 - El presidente Sebastián Piñera anunció, este martes, que el Transantiago cambiará de nombre para ser rebautizado como "Red Metropolitana ...

¿Fin del Transantiago? Ahora se va a llamar "Red" - Radio Imagina

<https://www.radioimagina.cl/2019/03/fin-del-transantiago-ahora-es-red/>

El presidente Sebastián Piñera anunció las medidas para reemplazar al sistema de Transantiago con nuevos buses Red".

Los cambios que traerá Red", el sistema que reemplazará al ... - Emol

<https://www.emol.com/.../Los-cambios-que-traera-Red-el-sistema-de-transportes-que-r...>

12 mar. 2019 - Los cambios que traerá Red", el sistema que reemplazará al Transantiago y la razón de su nombre. Una de las innovaciones apunta a la Bip!

Cambio de Transantiago a Red renovará el 70 % de los buses

<https://www.rockandpop.cl/2019/03/transantiago-red-renovara-buses/>

13 mar. 2019 - La modificación al sistema de transportes de la capital, que pasará después de más de una década como Transantiago a Red, cambiará los ...

Nuevo nombre para el Transantiago: Gobierno lo rebautizará como ...

<https://www.eldinamo.cl/.../nuevo-nombre-para-el-transantiago-gobierno-lo-rebautiza...>
11 mar. 2019 - El Gobierno decidió rebautizar el sistema de transporte público capitalino, conocido como Transantiago, el que ahora se llamará "Red".

Ministra Hutt: "No hemos gastado \$550 millones en un cambio de ...

<https://www.cnnchile.com/.../transantiago-red-550-millones-cambio-nombre-gloria-hu...>
12 mar. 2019 - "Red" es el nuevo nombre del sistema de transporte público ... a buses, al Metro y al sistema de ciclovías, en reemplazo del Transantiago.

De Transantiago a Red": Gobierno presenta nueva imagen del ...

<https://www.24horas.cl › Videos › Nacional>
"Transantiago cumplió, hizo lo que tenía que hacer y los contratos van terminado y lo estamos reemplazando por un sistema nuevo", afirmó la ministra de ...

Crece RED: tres nuevos recorridos, extensiones de servicios y ...

<https://www.publimetro.cl/cl/noticias/2019/.../red-transantiago-recorridos-trayectos.ht...>
hace 6 días - Las modificaciones del ex Transantiago beneficiarán a Cerrillos, Estación Central, Santiago, Pudahuel, Quilicura, Vitacura, Ñuñoa, La Reina y ...

¿Por qué el nuevo 'Transantiago' se llamará Red? - Sonar FM

<https://sonarfm.cl/sonarfm/noticias/por-que-el...transantiago-se...red/.../132757.html>
12 mar. 2019 - Este martes, el presidente Sebastián Piñera dio a conocer el nuevo sistema de transporte que operará en la Región Metropolitana y luego en ...

6.2.5. Consulta: comprar caliente camas

Calientacamás - Falabella.com

<https://www.falabella.com/falabella-cl/category/cat3046/Calientacamás>
Compra Calientacamás en Falabella.com: Calientacamás de 1 y 2 Plazas, para disfrutar de calor seguro todo el invierno. Encuentra el tuyo en Falabella.com. Scaldasofa Scaldasosno 1.5 ... · Calientacamás 2 Plazas, SG ...

Calientacamás | Sodimac.com

<https://www.sodimac.cl › Sodimac.com › Estufas › Calientacamás>
Encuentra todo lo que buscas en Calientacamás al mejor precio. Visita Sodimac.com y conoce nuestra gran variedad.

Calientacamás - ABCDIN

<https://www.abcdin.cl/tienda/es/abcdin/calientacamás>
Comprar. Ver detalle. Comparar Comparar. -8 %. Retiro en Tienda. Calientacamás Scaldasosno Sublime Maxi 2 Plazas. Internet \$66.990 Normal \$72.990.

Calientacamás | Ripley.com

<https://simple.ripley.cl/electro/climatizacion/calientacamás>

Descubre en ripley.com Calientacamás con increíbles precios y oportunidades exclusivas.

Ripley - Calientacamás - Ripley.com

<https://simple.ripley.cl/especial-electro/calientacamás>

CALIENTACAMA IMETEC SCALDASONNO MATRIMONIAL. \$44.990. En Promoción. KENDAL. CALIENTACAMA KENDAL INDIVIDUAL 1 PLAZA. \$16.990.

Calientacamás - Easy.cl

<https://www.easy.cl/tienda/categoria/calientacamás>

Encuentra en Easy.cl muebles, decoración, refrigeradores, herramientas y todo lo que necesitas para hacer de tu hogar el mejor lugar del mundo.

Calientacamás - Elige comodidad y calidez al dormir | Paris.cl

<https://www.paris.cl/dormitorio/ropa-cama/calientacamás/>

Olvidate del frío con estos prácticos calientacamás y disfruta de un sueño confortable sin preocuparte de las bajas temperaturas. Búscalos en Paris.

Calienta Camas en Mercado Libre Chile

<https://hogar.mercadolibre.cl> › Dormitorio › Ropa de Cama

Encuentra Calienta Camas en Mercado Libre Chile. Descubre la mejor forma de comprar online.

Recomendaciones para la compra y uso de calientacamás - SERNAC ...

<https://www.sernac.cl/portal/619/w3-article-5507.html>

26 jul. 2017 - Recomendaciones para la compra y uso de calientacamás Ver estudio completo Estudio del SERNAC comprobó que las 15 marcas ...

CALIENTACAMAS - Siegen

<https://siegen.cl/15-calientacamás>

Comprar. Ver ficha · imagen-product · Calientacama SG-5120D3. \$16.990. Comprar · Ver ficha · imagen-product. Calientacama SG-5120D4. \$16.990. Comprar.

CalientaCamás online. Calienta Camas en cic.cl | CIC

<https://cic.cl> › Ropa de Cama › Cubre Colchón › Calientacamás

Compra CalientaCamás en cic.cl. Este invierno no pases frío con los calientacamás de CIC. Conócelos online. ¡Visítanos!

Calientacamás - Beurer

www.beurer.cl/main/linea/71

Calientacamás. Calientacamás. Calor y ... CALIENTACAMAS TS19. 1 Plaza 1/2 ... Calientacamás UB40. 1 Plaza | ... Comprar Online · Beurer en Feria... Beurer ...

Calienta Cama Single CasaBella - Corona

<https://www.corona.cl> › Hogar › Linea Blanca › Climatización

li b Calienta Cama /b /li br li b Marca: Casa Bella /b /li br li b Single /b /li br li Material: 100 % poliéster /li br li Niveles de Seguridad.

Sernac explica en qué te debes fijar al elegir un calentacamas para ...

<https://www.biobiochile.cl/.../sernac-explica-en-que-te-debes-fijar-al-elegir-un-calienta...>

26 jul. 2017 - Sernac explica en qué te debes fijar al elegir un calentacamas para ... de recomendaciones al momento de comprar y usar un calentacamas, ...

Calientacamas Eléctrico Individual | Kendal Chile

<https://www.kendalchile.cl> › ... › Calientacamas › Calientacama Individual

Tengo un calentacamas Kendal Home y dejó de funcionar el mando regulador de encendido. ¿Dónde puedo comprar el repuesto? ¿Puedo instalarlo yo, o ...

calienta cama en Cuponatic

<https://www.cuponatic.com/buscador?q=calienta%20cama>

¿Cómo comprar? Publicar en Cuponatic · Nuestra Tienda · Cyberday. Contacto. ¿Necesitas Ayuda? 2 23071900. chile@cuponatic.com. Medios de pagos.

Comprar Calientacamas | Ofertas, precios y catálogos - Tiendeo.cl

<https://www.tiendeo.cl/ofertas/calientacamas>

Encuentra las mejores ofertas de Calientacamas y todas las promociones de tus tiendas favoritas. ¡Ahorra con Tiendeo!

Calientacamas en la mira: Conozca los peligros de usarlo. - Edifito

<https://www.edifito.com/calientacamas/>

26 jul. 2017 - Los calentacamas se usan entre el colchón y las sábanas, una vez ... 2017 que: "Lo principal es comprar sólo en el comercio establecido, ...

Sernac entregó recomendaciones para la compra y uso de calienta ...

<https://www.fortinmapocho.cl/.../sernac-entrego-recomendaciones-para-la-compra-y-u...>

26 jul. 2017 - Los calienta-camas constituyen un tipo de electrodoméstico de ... y recalcó que "es muy importante comprar en el comercio establecido, pues ...

Calientacama: ¿Durmiendo con el enemigo? - Chilevisión

www.chilevision.cl/matinal/reportajes/calientacama-durmiendo-con.../093225.html

27 may. 2015 - Conozca todos los mitos y verdades sobre el calentacamas en la ... con dolor de espaldas, no quiero pensar que es el calienta camas.

6.2.6. Consulta: 13c vivo online

En vivo Canal 13 | 13.cl

<https://www.13.cl/en-vivo>

En vivo. This video file cannot be played.(Error Code: 102630). En vivo. Teletrece Tarde. Hoy en el 13. Ver toda la programación. En vivoTeletrece Tarde. 15:25.

13C | 13.cl

<https://www.13.cl/c>

Hoy en 13C. Ver toda la programación. Ahora Síntoma Chilensis. 13:00. Reportajes de Selección. 13:30. Chilenautas. 14:30. Recomendando Chile. 15:30. Al Sur ... Programación · Plan V · Blog · Documentales

Inicio | 13.cl

<https://www.13.cl/>

Sitio web oficial de Canal 13, primer canal de transmisión chileno en transformarse en una casa editorial que traspasa sus contenidos a todas sus plataformas ...

En vivo Canal 13 | 13.cl

<https://www.13.cl/en-vivo?page=670>

En vivo. Este archivo de video no se puede reproducir.(Código de Error: 102630). En vivo. El Tiempo. Hoy en el 13. Ver toda la programación. En vivoEl Tiempo.

Canal 13 Cable En Vivo | TV Online Chile

tv-online.cl/canal-13-cable-en-vivo.php

Canal 13 Cable En Vivo, TV Online Chile: Sintoniza todas las senales abiertas de ... Tras su inicio, la senal de Canal 13C, conocida tambien como el Trece C, ...

Tele13 Radio 103.3 FM Video En Vivo | TV Online Chile

tv-online.cl/tele13-radio-103.3-fm-video-en-vivo.php

escuche a un canal de tv. que el gobierno va cambiar la hora de invierno en horario de verano, no cree señor ministro de energia que estadebe cambiarse en ...

Canal 13 Cable en Vivo | La Mejor |TV de Chile

www.lamejor.cl/canal13cable.html

Dejar un comentario del programa que estas viendo. Los comentarios publicados son de responsabilidad de quienes los emiten, no representan el ...

[EN VIVO] Sigue la transmisión de Tele13 Radio | Tele 13

<https://www.t13.cl/radio/en-vivo/videos/en-vivo-senal-video-de-tele13-radio>

13 jun. 2019 - Sigue las noticias y programación en la señal de video de Tele13 Radio 103.3 FM.

T13 Radio | Tele 13

<https://www.t13.cl/radio>

... Mundo · Nacional · Deportes · Tendencias · Reportajes T13. ESCUCHA AQUÍ. EN VIVO I TELE13 RADIO. Error loading player: No playable sources found ...

TV online - Radio online

cl.radioytv.com/tv.htm

Ve TV online. Señal en vivo. ... TVN, 24 horas, Mega, 13 C, Tele13, CNN Chile. Emol TV, La Tercera TV, La Cuarta TV, La Nación, CDF, BBC, Reino Unido.

Programación 13C, Lunes 24 de junio | Programación de TV en Chile ...

<https://mi.tv/cl/canales/canal-13-c>

Consulta todos los programas, series, películas etc. que echan hoy y en los próximos días en 13C.

Canal 13 - Señal Online - El Telón

<https://www.eltelon.com/television/canal-13>

Disfruta Canal 13 en vivo. Canal 13 online las 24hrs del día. Descubre cientos de canales de tv en vivo en nuestro sitio.

Tele13 Radio: Inicio

<https://www.tele13radio.cl/>

Conoce los programas y podcasts exclusivos de Tele13 Radio.

13C: Angélica Castro vuelve a la pantalla con el ... - La Opinión Online

<https://www.laopiniononline.cl/13c-angelica-castro-vuelve-a-la-pantalla-con-el-progra...>

3 oct. 2018 - "Discovery Chile", desde el jueves 4 de octubre solo por 13C. ... Me encanta la adrenalina y la sensación de en vivo , ya que lo que pasa en el ...

Obtener Canal 13: Microsoft Store es-CL

<https://www.microsoft.com/es-cl/p/canal-13/9wzdncrdjth7>

SofaScore LiveScore - Resultados en Vivo. Con calificación 5 de 5 ... Funciones. Ver señales online de Canal 13, 13C y REC TV, junto con la programación.

Hemisferio Derecho | Productora Hemisferio Derecho

www.hemisferioderecho.cl/

Serie Documental / Los Ángeles / 2019 / 4 Capítulos / Canal 13C. LIVE. 0 ... Sesiones en vivo de destacados músicos de la escena independiente chilena.

Como ver canales chilenos desde el extranjero? - ChW

www.chw.net › Foros › General › Foros de Relajo › Off Topic

29 ene. 2016 - La lata es que la señal de Canal 13c no esta disponible en el extranjero... Probé varias opciones Canal 13 Cable En Vivo | TV Online Chile ...

Comienza la nueva temporada de Soñadores por Canal 13 C

www.comprassustentables.cl/index.php?option=com_content...13-c...

6 ene. 2015 - La serie de televisión revelará semana a semana historias de emprendedores que realizan negocios con el Estado a través de la plataforma ...

Teatro Del Lago en Canal 13C | Teatro Del Lago

<https://www.teatrodelago.cl> › Noticias

... española - VIE 27 de DIC 2019; Close. Home » Noticias » Teatro Del Lago en Canal 13C. Noticia. Teatro Del Lago en Canal 13C. 13C - noviembre, 2017 ...

Londres 38 denuncia censura en serie "Una historia necesaria" de ...

<https://radio.uchile.cl/.../londres-38-denuncia-censura-en-serie-una-historia-necesaria-...>

14 sep. 2017 - Londres 38 denuncia censura en serie "Una historia necesaria" de Canal 13C. Según el espacio de memoria, durante la emisión de los ...

6.3. Resultados Yahoo

6.3.1. Consulta: profesores usm

Directorio de Personas USM

directorio.usm.cl

Profesores y Funcionarios Buscar : formatos aceptados: "nombre", "apellido", "nombre apellido" Titulados (listado de ex-alumnos que poseen dirección de email @titulados.usm.cl

Unidades Académicas y Docentes - usm.cl

www.usm.cl/universidad/unidades-academicas-docentes

Comunidad USM. Ir a Comunidad USM. Alumnos. Relaciones estudiantiles; Acceso SIGA; Aula; Portal de pagos; Webmail @alumnos; Webmail @sansano; Programas de intercambio; Federación de Estudiantes; Titulación 2017; Modelos; Profesores. Vicerrectoría Académica; Dirección General de Investigación, Innovación y Postgrado; Dirección General de ...

Comunidad USM - Universidad Técnica Federico Santa María

www.usm.cl/comunidad

Comunidad USM Conscientes de que nos enfrentamos a una sociedad cada vez más compleja y cambiante, con nuevas demandas, que imponen nuevos desafíos a la Universidad para ajustarse a esa realidad asumimos el compromiso de vincularnos con cada uno de nuestros miembros para proyectar la Universidad hacia el futuro.

electronica.usm.cl - Inicio · Departamento de Electrónica

www.electronica.usm.cl

Profesores, estudiantes, funcionarios, ex-alumnos y muchos más. Extensión: Conoce las distintas actividades extensión que tenemos para Toda la Comunidad. Accesos rápidos: Pregrado Postgrado Investigación Internacional Biblioteca USM Intranet ELO USM.cl Contáctanos. ... Valparaíso · +56 32 2654 000 · info@usm.cl.

Comunidad USM · Viña del Mar · Universidad Técnica ...

www.vinadelmar.usm.cl/comunidad-usm

Comunidad USM. Alumnos. Relaciones Estudiantiles Acceso SIGA Ramos Online Portal de Pagos ... Webmail @sansano Programas de intercambio Federación de Estudiantes Profesores. Vicerrectoría Académica Dirección General de Investigación y Postgrado Dirección General de Docencia Dirección de Estudios ...

Profesores, Técnico Universitario en Robótica y Mecatrónica

www.robotica.usm.cl/profesores

Profesor con 25 años de Experiencia como Docente en la Universidad Técnica Federico Santa María, especializado en Electrónica de Potencia, Control Automático y PLC. Jefe Proyecto Carrera T.U. Robótica y Mecatrónica

Profesores jornada parcial - industrias.usm.cl

www.industrias.usm.cl/quienes-somos/profesores-jp

Avenida España 1680, Edificio F1, Valparaíso · +56 (32) 2654272 comunicaciones.industrias@usm.cl Campus Santiago Vitacura. Avenida Santa María 6400, Vitacura, Santiago · +56 2 3202 8290 · comunicaciones.industrias@usm.cl. Sitio web creado por Dreyes.

Webmail UTFSM

webmail.usm.cl

@usm.cl Password. Mode. Language Cambiar clave. WebMail Antiguo Exchange Office365 Username. @usm.cl Password. Mode. Language. Cambiar clave. WebMail Antiguo Exchange Office365 ... WebMail Antiguo Exchange Office365 ...

sigas.usm.cl - Acceso

www.sigas.usm.cl

Para las Asignaturas: INTRODUCCION A LA FISICA, FISICA GENERAL I, MATEMATICA I, MATEMATICA II, QUIMICA Y SOCIEDAD y PROGRAMACION usted podrá elegir el paralelo (profesor) que desee responder de acuerdo a su asistencia en dichas asignaturas. (sólo para alumnos de primer año).

Plataforma Aula USM: Entrar al sitio

aula.usm.cl

¿Necesita ayuda para acceder a la plataforma de Aulas? » Para ingresar utilice su usuario y clave institucional » Si tiene problemas para acceder a la plataforma de Aulas, contáctese con la mesa de ayuda haciendo clic aquí.

Augusto Aninat S. - Ingeniería Comercial USM

comercial.usm.cl/quienes-somos/profesores/aninat-s

DEPARTAMENTO DE INGENIERÍA COMERCIAL Campus Casa Central Valparaíso. Avda. España 1680 Edificio F2, Valparaíso. +56 32 2654000. ingcomercial@usm.cl

Profesores, Técnico Universitario en Robótica y Mecatrónica

www.robotica.usm.cl/profesores

Profesor con 25 años de Experiencia como Docente en la Universidad Técnica Federico Santa María, especializado en Electrónica de Potencia, Control Automático y PLC. Jefe Proyecto Carrera T.U. Robótica y Mecatrónica

Profesores - mecanica.usm.cl

www.mecanica.usm.cl/10-profesores?start=15

Proyecto interno USM 25.15.11: ... Nuestros profesores y estudiantes de postgrado llevan a cabo además investigación fundamental y aplicada en los aspectos científicos de la disciplina, contribuyendo a la creación de nuevos conocimientos en los diversos campos de la espacialidad.

Profesores Jornada Completa - metalurgia.usm.cl

www.metalurgia.usm.cl/index.php/profesores/profesores-jornada-completa

Web del Departamento de Ingeniería Metalúrgica y de Materiales de la Universidad Técnica Federico Santa María

Profesores jornada parcial - industrias.usm.cl

www.industrias.usm.cl/quienes-somos/profesores-jp

Avenida España 1680, Edificio F1, Valparaíso · +56 (32) 2654272 comunicaciones.industrias@usm.cl Campus Santiago Vitacura. Avenida Santa María 6400, Vitacura, Santiago · +56 2 3202 8290 · comunicaciones.industrias@usm.cl. Sitio web creado por Dreyes.

Ignacio García S. - Ingeniería Comercial USM

comercial.usm.cl/quienes-somos/profesores/garcia-s

MBA, Universidad Politécnica de Cataluña, España. Magíster en Innovación Tecnológica y Emprendimiento, Universidad Técnica Federico Santa María, Chile.

Profesores - Universidad Técnica Federico Santa María

www.mecanica.usm.cl/quienes-somos/10-profesores

E-mail: rodrigo.barraza@usm.cl Área de especialización e interés Sistemas de procesos térmicos para aplicaciones de frío (criogenia) y calor utilizando fuentes de energías solares y convencionales.

Profesores Postgrado - electronica.usm.cl

www.electronica.usm.cl/profesores-postgrado

Profesores del Programa (Actualizado Junio 2019) Claustro de Profesores del Programa(CDTP*) (*Cuerpo de Directores de Tesis del Programa) [Magister] Área de Especialización del Programa Computadores Fernando Auat C. Grado (año) Doctor (2009) Institución otorgante Universidad Nacional de San Juan, Argentina Línea de Investigación Robótica Móvil Gonzalo Carvajal B. Grado (año) PhD (2013 ...

Webmail Profesores/Funcionarios :: Log in - mail.usm.cl

mail.usm.cl/horde

- Procedimiento de obtención (o cambio) de clave de CUENTA USM - ¿ Dónde solicito la creación de una cuenta de email ? - Información sobre Correo Electrónico en la USM - Listas de Email (Mailling Lists) - Configuración Acceso IMAP/POP de Mozilla Thunderbird o MS Outlook 2003.

Plataforma Aula USM: Entrar al sitio

aula.usm.cl

¿Necesita ayuda para acceder a la plataforma de Aulas? » Para ingresar utilice su usuario y clave institucional » Si tiene problemas para acceder a la plataforma de Aulas, contáctese con la mesa de ayuda haciendo clic aquí.

6.3.2. Consulta: eclipse solar 2019

Eclipse Solar Total 2 de Julio de 2019 - turismoastronomico.cl

www.turismoastronomico.cl/eclipse-solar-total-2-de-julio-de-2019.html

El 2 de julio de 2019 se producirá un eclipse solar total que será visible por una delgada franja del océano Pacífico y parte de América del Sur. La totalidad del eclipse comenzará al este de Nueva Zelanda en el Océano Pacífico para luego llegar a América del Sur. Los privilegiados esta vez serán los observadores de Chile y Argentina.

Eclipse solar del 2 de julio de 2019 - Wikipedia, la ...

es.wikipedia.org/wiki/Eclipse_solar_del_2_de_julio_de_2019

Eclipse solar del 2 de julio de 2019. El martes 2 de julio de 2019 se producirá un eclipse solar total que será visto en el océano Pacífico, Chile y al atardecer, en Argentina así como también en Uruguay.

TOUR ECLIPSE SOLAR 2019 La Serena, Higuera Chile 2 de ...

sunnytravel.cl/eclipse-solar

Eclipse Solar 2019 La Serena Tour en la higuera el 2 de julio de 2019 el mejor turismo ¡Únete a nosotros para la mejor experiencia de eclipse! RESERVA YA!

Eclipse solar 2019 en Chile - Astronomía en Chile y ...

www.chileestuyo.cl/eclipse-solar-2019-en-chile-lugares-para-el-astroturismo

Se realizarán observaciones de Eclipse desde el Cerro Patagual y el eclipse solar 2019 se verá en un 93,90 %. Muy próximo a la ciudad de Casablanca se encuentra el Observatorio Turístico Tokicura (4,6 km camino a Las Dichas). El eclipse en el Valle de Casablanca será parcial y se verá al 91,44 %.

Eclipse solar 2019: Esta será la duración en Chile del ...

www.ahoranoticias.cl/noticias/tendencias/260560-eclipse-solar-2019-en-chile...

Cada vez queda menos para el eclipse solar 2019 que podrá ser visto a la perfección en Chile, y las dudas sobre el fenómeno astronómico comienzan a surgir: desde por qué se ocasiona hasta cuánto durará.. Y si bien es un evento que pasa cada cierta cantidad de años en el país, los expertos han vuelto a hacer el llamado de lo peligroso que puede llegar a ver el eclipse sin estar ...

Calendario de Eclipses 2019 | Días de eclipses en 2019 ...

es.calenworld.com/calendario-lunar/calendario-de-eclipses-2019

Tipos de eclipses 2019. Existen dos tipos de eclipses diferentes, el eclipse lunar y el eclipse solar. Cada uno de ellos se diferencia en función de si está primero el Sol, la Tierra o la Luna, y así sabremos cuál se interpone y tapa la visión del otro. En 2019 se producirán 3 eclipses solares y 2 eclipses lunares.

Eclipse Solar Atacama 2019 - Atacama Eclipse Solar 2019

eclipsesolaratacama19.cl

El 2 de julio de 2019 se producirá un eclipse solar total que será visto en el océano Pacífi-

co, Chile y al atardecer, en Argentina así como también en Uruguay. Un eclipse solar ocurre cuando la Luna pasa entre la Tierra y el Sol, oscureciendo total o parcialmente la imagen del Sol para un espectador en la Tierra.

AURA - Chilean Eclipse 2019

www.eclipsesolar2019.cl

AURA - Chilean Eclipse 2019 La Asociación de Universidades para la Investigación en Astronomía (AURA) es un consorcio internacional de universidades e instituciones sin fines de lucro que operan instalaciones astronómicas de clase mundial .

Habrá eclipse solar el 2 de julio de 2019: en qué lugares ...

www.perfil.com/noticias/sociedad/eclipse-solar-2-julio-2019-lugares-donde-se...

Habrá eclipse solar el 2 de julio de 2019: en qué lugares se podrá ver La duración total del fenómeno será de casi cinco minutos y la Argentina será un escenario privilegiado para observarlo.

Solar Eclipse 2019 Chile

eclipsesolar2019.com

Les invitamos a participar de una de las más extraordinarias experiencias de este siglo, el esperado eclipse solar 2019 a realizarse en el hermoso y místico Valle del Elqui.

Eclipse Solar Atacama 2019 - Atacama Eclipse Solar 2019

eclipsesolaratacama19.cl

El 2 de julio de 2019 se producirá un eclipse solar total que será visto en el océano Pacífico, Chile y al atardecer, en Argentina así como también en Uruguay. Un eclipse solar ocurre cuando la Luna pasa entre la Tierra y el Sol, oscureciendo total o parcialmente la imagen del Sol para un espectador en la Tierra.

Eclipse Solar 2019 ¿Cómo influirá en los signos del ...

www.wemystic.com/es/eclipse-solar-2019

Al hablar de Eclipse Solar 2019, nos referimos a los tres eventos astrológicos de este tipo, que se registrarán durante los meses de enero, julio y diciembre de 2019, y cómo afectarán a los 12 signos del zodiaco.

Mapa del Eclipse - AURA - Chilean Eclipse 2019

www.eclipsesolar2019.cl/eclipse/mapa-del-eclipse

Credito: Observatorio Nacional Solar de EEUU, AURA y la Fundación Nacional de Ciencias de EE.UU. Usando la franja del Eclipse de NASA generado por Fred Espenak, Bessleian Elements para el Eclipse de Julio, 2019.

Se confirma eclipse solar total en Chile el 2019 ...

www.adnradio.cl/noticias/sociedad/se-confirma-eclipse-solar-total-en-chile-el-2019/...

El evento natural que llamó la atención de los estadounidenses durante esta semana ya tiene fecha, e incluso hora, en nuestro país: se viene un eclipse solar total el próximo 2 de Julio de

...

Eclipse solar 2020 - Araucanía | Eclipses Chile 2019 - 2020 v0

www.conicyt.cl/eclipseschile1/eclipse-solar-2020-araucania

El lunes 14 de diciembre de 2020, un eclipse total de Sol cruzará la Región de La Araucanía. En el resto del territorio nacional podrá observarse de manera parcial. El sur de Chile es la cuna de la cultura Mapuche, donde la observación de los astros y su influencia en la vida cotidiana, dieron origen a una rica cosmogonía.

Eclipse Solar Total 2019 EN VIVO: ¿Cuándo, dónde y a qué ...

rpp.pe/peru/actualidad/eclipse-solar-total-2019-en-vivo-cuando-donde-y-a-que...

Perú será uno de los países que podrá apreciar el eclipse solar del 2 de julio de 2019, pero de manera parcial. El último eclipse solar visto en nuestro país fue el del 2 de agosto de 2017

...

Habrá eclipse solar el 2 de julio de 2019: en qué lugares ...

www.perfil.com/noticias/sociedad/eclipse-solar-2-julio-2019-lugares-donde-se...

Habrá eclipse solar el 2 de julio de 2019: en qué lugares se podrá ver La duración total del fenómeno será de casi cinco minutos y la Argentina será un escenario privilegiado para observarlo.

Eclipse solar del 2 de julio de 2019 - Wikipedia, la ...

es.m.wikipedia.org/wiki/Eclipse_solar_del_2_de_julio_de_2019

Eclipse solar del 2 de julio de 2019. El martes 2 de julio de 2019 se producirá un eclipse solar total que será visto en el océano Pacífico, Chile y al atardecer, en Argentina así como también en Uruguay.

Solar Eclipse 2019 Chile

eclipsesolar2019.com

Les invitamos a participar de una de las más extraordinarias experiencias de este siglo, el esperado eclipse solar 2019 a realizarse en el hermoso y místico Valle del Elqui.

El próximo eclipse solar total será en 2019 y podrá verse ...

www.lanacion.com.ar/sociedad/el-proximo-eclipse-solar-total-sera-en-2019-y...

El próximo eclipse solar total será en 2019 y podrá verse en la Argentina - Se verá en su totalidad en el centro de San Juan, sur de La Rioja, norte de San Luis, centro de Córdoba, sur de ...

6.3.3. Consulta: radio carolina

Radio Carolina

www.carolina.cl

A pesar de ganarle a Tailandia por 2-0, le faltó un gol más para clasificar a octavos de final.

Radio Carolina online - Señal en vivo - 99.3 MHz FM ...

<https://onlineradiobox.com/cl/carolina>

Radio Carolina 99.3 MHz FM, Santiago de Chile, Chile - escuchar radio online gratis en On-lineRadioBox.com. Esta página web utiliza cookies. Al continuar utilizando esta página web, estás de acuerdo con nuestras políticas respecto al uso de cookies.

RADIO CAROLINA - ON LINE - SEÑAL EN VIVO - [http://www ...](http://www...)

www.talcaguia.cl/radio_online.php?id=19

radio carolina en vivo señal online gratis chile en directo - on line señal en vivo en directo
<http://www.carolina.cl/>

Radio Carolina, 99.3 FM, Santiago, Chile | Free Internet ...

<https://tunein.com/radio/Radio-Carolina-993-s25096>

Radio Carolina - Hiperkinética y de permanente buen humor, Carolina trae lo mejor de los sonidos de moda en una selección que incluye música nacional, Electronica, Pop y Reggaetón.

Radio Carolina - Wikipedia, la enciclopedia libre

https://es.wikipedia.org/wiki/Radio_Carolina

A comienzos de 2009 Radio Carolina abandona la frecuencia 100.3 de San Antonio, frecuencia ocupada hasta 2017 por Radio Beethoven, provocando molestia entre sus auditores de esa zona. Desde la radio no dieron ningún tipo de explicaciones sobre su salida.

Eslogan: La Más Prendida, Yo Escucho Carolina

Radios Online Chile, emisoras chilenas, radios en vivo

<https://emisora.cl> ¿Buscas una radio online de Chile? En nuestra web encontrarás tu radio en vivo favorita, ¡con toda la música online a un solo clic! ¡Fácil, rápido y ameno!

Radio Carolina - Señal Online - El Telón - Radio y ...Traducir esta página

www.eltelon.com/radios/carolina

Disfruta Radio Carolina online las 24hrs del día. Descubre cientos de radios en vivo en nuestro sitio.

Mi Pagina, Portal de Sitios Chilenos

mipagina.cl

Miles de sitios chilenos útiles , clasificados en categorías. Todos los Diarios y radios de Chile. .Msn Messenger via web. Escuchar todas las radios online de Chile . Traductor. Medidor de conexion

Mi Pagina, Portal de Sitios Chilenos

mipagina.cl

Miles de sitios chilenos útiles , clasificados en categorías. Todos los Diarios y radios de Chile. .Msn Messenger via web. Escuchar todas las radios online de Chile . Traductor. Medidor de conexion

Radio Carolina online - Escuchar Radio on line

<https://radiochilena.net/radio-carolina>

Radio Carolina. Radio Carolina es una de las estaciones preferidas por la juventud chilena,

ya que incorpora música urbana actual tales como reggaetón, hip-hop, EDM, dance y hasta pop.

Radio Carolina - guiapenquista.cl

www.guiapenquista.cl/radios-chilenas/radio-carolina.php

Radio Carolina online señal en vivo frecuencia 99.3 Fm por internet para Chile y el Mundo

Radio Karolina - Lagana i fina

<https://karolina.rs>

Radio Karolina je prisutan u srpskom medijskom prostoru od 2013. godine. Najpre je bila pokrivena samo teritorija Beograda, da bi u međuvremenu i Novi Sad i Niš dobili svoje frekvencije. Radio je koncipiran tako da se slušaocima, ali i tržištu uopšte, ponudi program koji je jedinstven u Srbiji.

Radio Caroline Traducir esta página

www.radiocaroline.co.uk

Radio Caroline, the former offshore radio station, now sending multiple music streams on the internet, also local full time DAB and regional AM on special weekends

Radio Carolina - Señal Online - eltelon.com

www.eltelon.com/radios/carolina?in=chile

Disfruta Radio Carolina online las 24hrs del día. Descubre cientos de radios en vivo en nuestro sitio.

RADIO CAROLINA - Home | Facebook Traducir esta página

<https://www.facebook.com/RADIOCAROLINA>

RADIO CAROLINA, Ñuñoa. 1,158,790 likes · 79,641 talking about this. Radio Carolina! 🎧 La Más Prendida 🎧

Uživo Radio Karolina Beograd, Karolina radio streaming ...

www.uzivoradio.com/karolina-beograd.html

Radio Karolina lagana i fina iz Beograda uživo. Slušajte Karolina radio uživo, internet radio stream. We use cookies to ensure that we give you the best experience. If you continue using this website, we'll assume that you are OK about that.

Radio Carolina - En vivo en línea Radio

<https://www.liveonlineradio.net/es/chile/radio-carolina.htm>

Radio Carolina basado en La Serena, Chile es uno de la estación de la música popular. Estación de Radio Carolina música en streaming y Programms tanto en el aire y en línea. Originalmente es una música española, Adulto Contemporáneo, Top canal de radio 40 / ...

Daddy Yankee - Dura (Video Oficial) - YouTube

6.3.4. Consulta: red transantiago

Red Metropolitana de Movilidad

www.red.cl

Con nuevos servicios y extensiones, Ministerio de Transportes estrena malla de recorridos que beneficiará a nueve comunas de Santiago. Los cambios, que debutan el domingo, contemplan tres nuevos recorridos, dos ampliaciones y once ajustes que darán más conectividad a Cerrillos, Estación Central, Santiago, Pudahuel, Quilicura, Vitacura, Ñuñoa, La Reina y Peñalolén.

Las características de Red que lo hacen distinto al ...

www.fayerwayer.com/2019/03/red-transantiago-caracteristicas

Red ya comienza su implementación y lo que queda de acá en más es descubrir exactamente qué diferencias existirán con el famoso y tan vilipendiado Transantiago. Según lo que han prometido las autoridades, este sistema se fundamenta en contratos más cortos y flexibles, con empresas de buses ...

De Transantiago a Red: ¿Solución real o mera estrategia ...

radio.uchile.cl/2019/03/12/de-transantiago-a-red-solucion-real-o-mera...

El presidente Sebastián Piñera anunció, este martes, que el Transantiago cambiará de nombre para ser rebautizado como “Red Metropolitana de Movilidad”. Según comentó el mandatario, la nueva imagen se debe a que será una “red” que conectará los buses, el Metro, Metrotren y bicicletas ...

Mapas de recorridos | Red Metropolitana de Movilidad

www.red.cl/mapas-y-recorridos/mapas-de-recorridos

A continuación encontrarás el mapa de las 32 comunas de la Región Metropolitana, más San Bernardo y Puente Alto, junto con las rutas habituales de todos los recorridos de buses de las empresas concesionarias que operan en estas comunas.

Experto prevé alza de tarifas de Red (ex Transantiago ...

www.cooperativa.cl/noticias/pais/transportes/transantiago/experto-preve-alza...

Nuevos recorridos del Transantiago beneficiarán a 700 mil personas Gobierno dio el primer paso formal en cambio de Transantiago a Red Santiago tendrá el primer corredor 100 % eléctrico de ...

Dale dale RED: Así es el nuevo sistema de transporte que ...

www.fayerwayer.com/2019/03/red-transantiago-transporte

RED, ese es el nombre del nuevo sistema de transporte que reemplazará al Transantiago y que este martes 12 de marzo presentó el presidente Sebastián Piñera y la ministra Gloria Hutt. Hola Bienvenidos. Te invitamos a conocer @Red_Movilidad el nuevo nombre del Sistema del Transporte ...

Nuevo sistema de transporte público: las diferencias entre ...

www.24horas.cl/nacional/nuevo-sistema-de-transporte-publico-las-diferencias...

Este martes, el Presidente Sebastián Piñera junto a la ministra de Transportes, Gloria Hutt, presentó Red”, el nuevo sistema de transporte público que reemplaza al Transantiago y anunció la ...

Cambio de Transantiago a Red renovará el 70 % de los buses

www.rockandpop.cl/2019/03/transantiago-red-renovara-buses

13 marzo, 2019 Cambio de Transantiago a Red renovará el 70 % de los buses La modificación al sistema de transportes de la capital, que pasará después de más de una década como Transantiago a Red, cambiará los buses.

Las reacciones que ha dejado Red”, el reemplazo del ...

www.fmdos.cl/noticias/reacciones-red-sistema-reemplaza-transantiago

Este martes el presidente Sebastián Piñera presentó el nuevo sistema de transporte «Red» (Red Metropolitana de Movilidad), que viene a reemplazar al Transantiago tras 11 años de su funcionamiento.. Fueron 550 millones de pesos invertidos en la asesoría comunicacional que buscó darle un nuevo rostro al transporte público de la capital, según informa «La Tercera».

Red (Transantiago) | Interior Bus de 2 pisos Wrightbus ...

www.youtube.com/watch?v=3_JHyuaaFPw

#RedMejorTransporte #Wrightbus #BusesVule Grabación de exterior e interior de un bus a prueba del transporte público de Santiago de Chile, unidad Wrightbus Streetdeck 10.6 Euro 6. Este bus está ...

De Simpsons y cumbias: Las reacciones por la llegada de ...

mouse.latercera.com/red-transantiago-memes

12 MAR 2019 / Actualidad De Simpsons y cumbias: Las reacciones por la llegada de Red y el adiós al Transantiago. No pocos clamaron que, ante el nuevo sistema de transporte capitalino, corearán: “Este es el color de mi locomoción”.

Cambio de Transantiago a Red renovará el 70 % de los buses

www.rockandpop.cl/2019/03/transantiago-red-renovara-buses

Cambio de Transantiago a Red renovará el 70 % de los buses La modificación al sistema de transportes de la capital, que pasará después de más de una década como Transantiago a Red, cambiará los buses. Por Nicolás Noli.

En qué consistirá Red”: El nuevo sistema de transporte ...

www.ahoranoticias.cl/noticias/nacional/253643-red-cambio-transantiago-modifica...

Nacional En qué consistirá Red”: El nuevo sistema de transporte capitalino que reemplazará al Transantiago Entre las novedades destaca el cambio de colores y estándares para operadores.

Las reacciones que ha dejado Red”, el reemplazo del ...

www.fmdos.cl/noticias/reacciones-red-sistema-reemplaza-transantiago

Este martes el presidente Sebastián Piñera presentó el nuevo sistema de transporte «Red» (Red Metropolitana de Movilidad), que viene a reemplazar al Transantiago tras 11 años de su

funcionamiento.. Fueron 550 millones de pesos invertidos en la asesoría comunicacional que buscó darle un nuevo rostro al transporte público de la capital, según informa «La Tercera».

Nuevo nombre para el Transantiago: Gobierno lo rebautizará ...

www.eldinamo.cl/nacional/2019/03/11/nuevo-nombre-para-el-transantiago-gobierno...

Nacional Nuevo nombre para el Transantiago: Gobierno lo rebautizará como “Red” La Monda pagó 550 millones de pesos a la empresa Dittborn y Unzueta, a través de Chilecompra, por la ...

Transantiago tendrá un nuevo nombre: se llamará Red ...

www.biobiochile.cl/noticias/nacional/region-metropolitana/2019/03/12/aseguran...

El Gobierno del presidente Sebastián Piñera ha prometido en reiteradas ocasiones una transformación completa al Transantiago, sistema de transportes que prometía mejorar la forma de ...

Red: el nuevo nombre que tendrá el Transantiago - t13.cl

www.t13.cl/noticia/nacional/red-nuevo-nombre-tendra-transantiago

Red”: el nuevo nombre que tendrá el Transantiago. Presidente Piñera realizará este martes una ceremonia para presentar el “cambio de identidad” que tendrá el transporte público de la Región Metropolitana.

De Transantiago a Red: Memes hacen humor con el cambio de ...

www.guioteca.com/fotografia/de-transantiago-a-red-memes-hacen-humor-con-el...

“Red podrá ser replicado a diferentes ciudades del país, adaptándose a las realidades locales”, finalizó el Presidente Sebastián Piñera respecto a un anuncio que generó diversas reacciones en las redes sociales, entre ellas los tradicionales memes que hicieron humor con el cambio del Transantiago a “Red”. Aquí, algunas reacciones:

[VIDEO] Red: El fin del Transantiago | Tele 13

www.t13.cl/videos/nacional/video-red-fin-del-transantiago

VIDEO RED: El fin del Transantiago. Tras 12 años de operación, el nombre “Transantiago” está viviendo sus últimos días. El Gobierno presentó hoy el nuevo sistema de transporte Público que ahora se llamará RED.

De Transantiago a Red – Noticias UAI

noticias.uai.cl/de-transantiago-a-red

El presidente Sebastián Piñera anunció que el Transantiago cambiará de nombre a “Red Metropolitana de Movilidad”, y comentó que la nueva imagen se debe a que será una “red” que conectará los buses, el metro, metrotren y bicicletas, y que el cambio llegará con la puesta en marcha de 200 buses eléctricos y 490 ecológicos, lo que significaría un mejor estándar de servicio en ...

6.3.5. Consulta: comprar caliente camas

Caliente camas | Hites

www.hites.com/electro-hogar/climatizacion/calientacamam

Compra Caliente camas en nuestras tiendas Hites. En Hites.com podrás equipar tu casa en electrohogar con los mejores Refrigeradores, Lavadoras, Secadoras, Cocina, Cuidado Personal, Electrodomésticos y Climatización.

Ofertas en Todos las Calientacamam | Abcdin.cl - Tienda Online

www.abcdin.cl/tienda/es/abcdin/calientacamam

Calientacamam en ofertas a precios imperdibles, calienta tu cama en Abcdin.cl | La felicidad cuesta menos Aquí! Cerrar Comparar Agrega otro Producto ... Comprar Ver detalle Comparar Comparar Retiro en Tienda. Calientacamam Scaldasonno Matrimonial Internet \$32.990 Normal \$44.990

Calientacamam | Sodimac.com

www.sodimac.cl/sodimac-cl/category/scat299483/Calientacamam

Encuentra todo lo que buscas en Calientacamam al mejor precio. Visita Sodimac.com y conoce nuestra gran variedad.

Guía de Compra de Calientacamam Eléctricos (2019)

elvergelweb.com/guia-de-compra-de-calientacamam-electricos

Guía de Compra de Calientacamam Eléctricos. Bienvenida/o, a otra guía de compra de ElVergelWeb.com, en esta ocasión hablaremos de los Calientacamam Eléctricos. En este artículo te detallaré las características más importantes que debes tener en cuenta para poder tomar una buena decisión de compra.

Comprar calientacamam y cubrecolchones baratos - Orbegozo ...

www.electrocosto.com/calientacamam

¿Qué calientacamam eléctrico comprar? Consejos y Recomendaciones. Las mejores calientacamam eléctricos. Las mejores ofertas en calientacamam eléctricos las vas a encontrar en Electrocosto. En invierno si quieres proteger tu cama del frío y la humedad el cubrecolchones o calientacamam puede ser la solución.

caliente camas en abcdin.cl

www.abcdin.cl/tienda/search/caliente+camas

caliente camas 1.251 ... Comprar Ver detalle Comparar Comparar Despacho a Domicilio. Diván Cama Rosen Ergo T 1,5 Plazas Internet \$239.990 Normal \$314.990 Comprar Ver detalle Comparar Comparar Despacho a Domicilio. Diván Rosen Ergo-t 1,5 Plazas + Respaldo + Vela-dor Tabor ...

Comprar Calientacamam | Ofertas, precios y catálogos

www.tiendeo.cl/ofertas/calientacamam

Comprar calientacamam - Ofertas y promociones Favorito . Ofertas. Ofertas (5) Catálogos (5) Cupones (0) Catálogos de Calientacamam y de Hogar y muebles en tu ciudad. Ha habido un

problema inténtelo más tarde. ¡Añade Paris a tus Favoritos! Catálogo Paris En Tienda Válido 11 días.

Calientacamás | Ripley.com

simple.ripley.cl/electro/climatizacion/calientacamás

Comprar en Ripley.com es 100 % seguro. Comercial ECCSA S.A., Rut N° 83.382.700-6, domiciliada en Avenida Cerro Colorado 5240, Torre del Parque 1 Piso 11, comuna Las Condes, Santiago de Chile. Precios publicados aplican exclusivamente a Internet y Venta Telefónica, estos pueden llegar a ser distintos en tiendas.

Comprar Calientacamás en Curicó - Ofertas y tiendas - Ofertia

www.ofertia.cl/curico/ofertas/calientacamás

¡Encuétralas en Ofertia! Todas las tiendas en Curicó con catálogo de Calientacamás. Accede sencillamente a catálogos, ofertas y folletos de las tiendas a tu alrededor. Producto, marca o nombre de la tienda. Ciudad. Buscar. ... Tiendas donde comprar Calientacamás en Curicó ...

Calienta Camas en Mercado Libre Chile

hogar.mercadolibre.cl/dormitorio/ropa-cama/calienta-camas

Encuentra Calienta Camas en Mercado Libre Chile. Descubre la mejor forma de comprar online.

Guía de Compra de Calientacamás Eléctricos (2019)

elvergelweb.com/guia-de-compra-de-calientacamás-electricos

Guía de Compra de Calientacamás Eléctricos. Bienvenida/o, a otra guía de compra de ElVerGelWeb.com, en esta ocasión hablaremos de los Calientacamás Eléctricos. En este artículo te detallaré las características más importantes que debes tener en cuenta para poder tomar una buena decisión de compra.

Mejor Calientacamás Eléctrico junio 2019 -Comparativa

consejosparaahorrar.net/hogar-y-cocina/manta-electrica/mejores-calientacamás...

A continuación puede leer nuestra guía de compra para conocer sus características, tipos y saber comprar las mejores calientacamás térmicas en base a criterios de calidad, precio, ofertas e información correcta para el consumidor, pero de todas formas, si no quiere usar ese tiempo, le dejamos la lista de las mantas eléctricas para camas mas vendidas del año en Amazon!.

Ripley - CALIENTA CAMA CIC KING

simple.ripley.cl/calienta-cama-cic-king-2000346214053p

CALIENTACAMA CIC KING. Para dormir plácidamente sin preocuparte del frío, este calienta-cama de CIC es la solución ideal. Con un consumo de sólo 60 watts por lado, está confeccionado 100 % en poliéster de alta calidad con 3 capas de telas.

calienta camas en abcdin.cl

www.abcdin.cl/tienda/search/calienta+camas

calienta camas 1.251 ... Comprar Ver detalle Comparar Comparar Despacho a Domicilio. Di-

ván Cama Rosen Ergo T 1,5 Plazas Internet \$239.990 Normal \$314.990 Comprar Ver detalle Comparar Comparar Despacho a Domicilio. Diván Rosen Ergo-t 1,5 Plazas + Respaldo + Vela-dor Tabor ...

Comprar un Calientacamás - La Mejor Guía de Compra para el ...

tumantaelectrica.com/calientacamás-más-vendidos

Comparativa de los ocho calientacamás más vendidos. En esta tabla comparativa actualizada en Enero del 2018 vas a poder comparar fácilmente los precios y las características principales de los calientacamás con mejor relación calidad-precio. Además son los calientacamás más vendidos en Amazon: comprueba las valoraciones de otros clientes que han comprado el producto.

Calientacamás - Elige comodidad y calidez al dormir | Paris.cl

www.paris.cl/dormitorio/ropa-cama/calientacamás

Olvídate del frío con estos prácticos calienta camas y disfruta de un sueño confortable sin preocuparte de las bajas temperaturas. Búscalos en Paris.

Comprar calientacamás en San Joaquín - Ofertas ...

www.tiendeo.cl/san-joaquin/ofertas/calientacamás

Comprar calientacamás en San Joaquín - Ofertas, promociones y cupones Seguir . Ver oferta . 18 días más. Especial belleza . 1. Bolsa de agua caliente. Ver oferta . 4 días más. Especial día del padre . 1. Calientapiés. Ver oferta . 18 días más. Aprovecha tus puntos y ahorra . 1. Manta eléctrica. Ver oferta .

Compra Calientacamás Matrimonial online | Linio Chile

www.linio.cl/p/calientacamás-matrimonial-qf3awm

Compra Calientacamás Matrimonial online Encuentra los mejores productos Calefacción Kendal en Linio Chile

CALIENTACAMAS - Siegen

siegen.cl/15-calientacamás

Calientacama SG-5121D1 \$29.990 Comprar

CALIENTACAMAS - Siegen

siegen.cl/15-calientacamás?p=2

Calientacama SSG-5501 \$99.990 Comprar

6.3.6. Consulta: 13c vivo online

En vivo Canal 13 | 13.cl

www.13.cl/en-vivo

Sitio web oficial de Canal 13, primer canal de transmisión chileno en transformarse en una casa editorial que traspasa sus contenidos a todas sus plataformas como Internet, radio, cable y eventos multitudinarios.

13C | 13.cl

www.13.cl/c

Inés Matte Urrejola 0848, Santiago, Chile - Fono (562) 2 251 4000 © Todos los derechos reservados. 13.cl

Canal 13 Cable En Vivo | TV Online Chile

tv-online.cl/canal-13-cable-en-vivo.php

Canal 13 Cable transmite de forma ininterrumpida las 24 horas del día y gracias a tv-online.cl podras ver su senal en directo desde tu telefono, computador o Tablet. No esperes mas y comienza a disfrutar de todo el contenido de Canal 13C hoy mismo y Ve la cultura con otros ojos!

Canal 13 C Chile en vivo, Online Teleame Directos TV

directostv.teleame.com/canal13cchileenvivo

Canal 13 C Chile en vivo, Online, en directo. 13C (anteriormente Señal 3 y Canal 13 Cable) es un canal de televisión por cable chileno, de propiedad y operado por Canal 13 S.A.

Canal 13 en vivo - (El trece) Mirá todos los canales de TV ...

www.televisionparatodos.tv/canal-13-en-vivo-online

Canal 13 en vivo, actualmente conocida por su nombre comercial El Trece, es una cadena de televisión argentina que transmite desde la Ciudad Autónoma de Buenos Aires, siendo la emisora con mayor cobertura nacional. Es la encargada de proporcionar contenidos de programación a diferentes canales de televisión en el interior del país, a través de su señal satelital.

CANAL 13 | ON LINE SEÑAL EN VIVO EN DIRECTO GRATIS

www.talcaguia.cl/tv_online.php?id=8

canal 13 | on line seÑal en vivo gratis. canal 13 | on line seÑal en vivo gratis ... además, posee desde 1995 una segunda seÑal en el cable, llamada 13c (antes seÑal 3), con programación cultural, y tiene a su cargo teleduc, creado en 1977, que es el centro de educación a distancia de la universidad catÓlica.

TV online, Canales de Television de Chile, Señal en vivo

www.refnet.cl/tvonline.html

TV online, Señal en vivo Canales de Television de Chile, Windows Media, Flash Player

ON LINE SEÑAL EN VIVO EN DIRECTO GRATIS - talcaguia.cl

www.talcaguia.cl/tv_online.php?id=181

Eleodoro Rodríguez Matte, director ejecutivo en 1995, decide internacionalizar la señal Canal 13, la estación más vista a nivel nacional. Para ello inicia la creación de dos frecuencias paralelas a la señal abierta: señal internacional, vía satélite; y la señal 3, vía cable.

Señal Online - Chilevisión

www.chilevision.cl/senal-online

Disfruta del contenido de ChileVisión en nuestra señal online. Somos el canal oficial de la

Selección Chilena de Fútbol y #VamosContigo

SEÑAL EN VIVO MEGA | Mega.cl

www.mega.cl/senal-en-vivo

Te recomendamos ¡Felipe Lasso se casa hoy y así va la previa de la ceremonia! [VIDEO] Cami Gallardo da contundente respuesta a "haters.^{en} nuevo y sensual videoclip

envivo.13.cl

envivo.13.cl/13.php

envivo.13.cl ... 13 13

ON LINE SEÑAL EN VIVO EN DIRECTO GRATIS - talcagua.cl

www.talcagua.cl/tv_online.php?id=181

Eleodoro Rodríguez Matte, director ejecutivo en 1995, decide internacionalizar la señal Canal 13, la estación más vista a nivel nacional. Para ello inicia la creación de dos frecuencias paralelas a la señal abierta: señal internacional, vía satélite; y la señal 3, vía cable.

Señal Online - Chilevisión

www.chilevision.cl/senal-online

Disfruta del contenido de ChileVisión en nuestra señal online. Somos el canal oficial de la Selección Chilena de Fútbol y #VamosContigo

Canal 13 En Vivo | Tv Chilena Online

chilena-tv.blogspot.com/2016/10/canal-13-chile-en-vivo.html

Canal 13 en vivo es propiedad y creada por ingenieros de La Universidad Católica inició oficialmente las transmisiones en la frecuencia 2. Sin duda, el hito más importante y que dio el verdadero puntapié inicial de este nuevo medio en el país, fue la transmisión del Campeonato Mundial de Fútbol en 1962, realizado en Chile

SEÑAL EN VIVO MEGA | Mega.cl

www.mega.cl/senal-en-vivo

Te recomendamos ¡Felipe Lasso se casa hoy y así va la previa de la ceremonia! [VIDEO] Cami Gallardo da contundente respuesta a "haters.^{en} nuevo y sensual videoclip

Radios Online Chile, emisoras chilenas, radios en vivo

emisora.cl

¿Buscas una radio online de Chile? En nuestra web encontrarás tu radio en vivo favorita, ¡con toda la música online a un solo clic! ¡Fácil, rápido y ameno!

VER CANAL 13 EN VIVO ONLINE GRATIS POR INTERNET 📺 TV ...

www.youtube.com/watch?v=imaagK7pWIE

Después de muchas idas y vueltas, Vuelve el show: Marcelo Tinelli firmó contrato con El Trece 13 en vivo por internet television argentina señal online canal abierto en HD.

Ver UCV en vivo - UCV online

vertelevisionenvivo.blogspot.com/2010/08/ucv-television-de-chile-en-vivo.html

Disfruta del enlace a la señal en vivo y en directo de UCV, aquí en Ver Television en Vivo, pronto incluiremos la programación actualizada de UCV a su alcance, Enlazar este post: Entrada más reciente Entrada antigua Página principal >MenuCountry. Checkcountry Delete country EditCountry ...

Televisión Nacional de Chile

www.tvn.cl

Sitio web de Televisión Nacional de Chile, TVN. Somos el canal de Chile y nuestra programación ofrece programas de noticias de Chile y el mundo, reportajes, debate, servicio, telenovelas, series, películas, cultura, espectáculos y deporte.

TV en vivo - Canales de Televisión por internet

www.onlinetv.planetfools.com

Los mejores eventos y fútbol en vivo de la televisión están en online TV. Noticias, telenovelas, entretenimiento, espectáculos, deportes y fútbol en vivo 7 días a la semana, 24 horas al día. Contamos con más de 200 canales en español donde podrás disfrutar del mejor entretenimiento gratis. TV de México, TV de USA, TV de América ...

6.4. Resultados Bing

6.4.1. Consulta: profesores usm

Directorio de Personas USM

directorio.usm.cl

Profesores y Funcionarios Buscar : formatos aceptados: "nombre", "apellido", "nombre apellido" Titulados (listado de ex-alumnos que poseen dirección de email @titulados.usm.cl

Comunidad USM - Universidad Técnica Federico Santa María

www.usm.cl/comunidad

Comunidad USM Conscientes de que nos enfrentamos a una sociedad cada vez más compleja y cambiante, con nuevas demandas, que imponen nuevos desafíos a la Universidad para ajustarse a esa realidad asumimos el compromiso de vincularnos con cada uno de nuestros miembros para proyectar la Universidad hacia el futuro.

Unidades Académicas y Docentes - usm.cl

<https://www.usm.cl/universidad/unidades-academicas-docentes>

Comunidad USM. Ir a Comunidad USM. Alumnos. Relaciones estudiantiles; Acceso SIGA; Aula; Portal de pagos; Webmail @alumnos; Webmail @sansano; Programas de intercambio; Federación de Estudiantes; Titulación 2017; Modelos; Profesores. Vicerrectoría Académica; Dirección General de Investigación, Innovación y Postgrado; Dirección General de ...

Comunidad USM · Viña del Mar · Universidad Técnica ...

www.vinadelmar.usm.cl/comunidad-usm

Comunidad USM. Alumnos. Relaciones Estudiantiles Acceso SIGA Ramos Online Portal de Pagos ... Webmail @sansano Programas de intercambio Federación de Estudiantes Profesores. Vicerrectoría Académica Dirección General de Investigación y Postgrado Dirección General ...

electronica.usm.cl - Inicio • Departamento de Electrónica

www.electronica.usm.cl

Profesores, estudiantes, funcionarios, ex-alumnos y muchos más. Extensión: Conoce las distintas actividades extensión que tenemos para Toda la Comunidad. Accesos rápidos: Pregrado Postgrado Investigación Internacional Biblioteca USM Intranet ELO USM.cl Contáctanos. ... Valparaíso • +56 32 2654 000 • info@usm.cl.

Webmail UTFSMTraducir esta página

<https://webmail.usm.cl>

@usm.cl Password. Mode. Language Cambiar clave. WebMail Antiguo Exchange Office365 Username. @usm.cl Password. Mode. Language. Cambiar clave. WebMail Antiguo Exchange Office365 ... WebMail Antiguo Exchange Office365 ...

Inicio • Defider • Universidad Técnica Federico Santa María

defider.usm.cl

USM inaugura nuevo Gimnasio 3 en Campus Casa Central Conoce más. USM reconoce y premia a los estudiantes y profesores destacados durante 2018 Conoce más. DEFIDER premia a los mejores de este 2018 en el Campus Santiago

Plataforma Aula USM: Entrar al sitio

aula.usm.cl

¿Necesita ayuda para acceder a la plataforma de Aulas? » Para ingresar utilice su usuario y clave institucional » Si tiene problemas para acceder a la plataforma de Aulas, contáctese con la mesa de ayuda haciendo clic aquí.

sig.usm.cl - Acceso

www.sig.usm.cl

La finalidad principal de la encuesta es contribuir al perfeccionamiento de la docencia y al proceso enseñanza aprendizaje; los datos recogidos son importantes para la Universidad, los profesores y los propios alumnos. Se agradece la participación seria, responsable y oportuna asegurándose el anonimato y la debida consideración de sus ...

USM rinde homenaje al destacado profesor Walter Grote ...

<https://noticias.usm.cl/2019/05/28/usm-rinde-homenaje-al-destacado-profesor-walter-grote>

El académico perteneciente al Departamento de Electrónica, fue reconocido por su trayectoria de más 40 años y sus importantes aportes a la Institución. En una emotiva ceremonia en la que participaron profesores, estudiantes y funcionarios, la Universidad Técnica Federico Santa María rindió un merecido homenaje, por su más de cuarenta años de trayectoria, al [...]

Profesores jornada parcial - industrias.usm.cl

www.industrias.usm.cl/quienes-somos/profesores-jp

Avenida España 1680, Edificio F1, Valparaíso · +56 (32) 2654272 comunicaciones.industrias@usm.cl Campus Santiago Vitacura. Avenida Santa María 6400, Vitacura, Santiago · +56 2 3202 8290 · comunicaciones.industrias@usm.cl. Sitio web creado por Dreyes.

Profesores - Universidad Técnica Federico Santa ...Traducir esta página

www.mecanica.usm.cl/quienes-somos/10-profesores

E-mail: rodrigo.barraza@usm.cl Área de especialización e interés Sistemas de procesos térmicos para aplicaciones de frío (criogenia) y calor utilizando fuentes de energías solares y convencionales.

Profesores, Técnico Universitario en Robótica y Mecatrónica

www.robotica.usm.cl/profesores

MSc. Mechatronics, University of Newcastle, England. Ingeniero Civil Mecánico, Universidad de Concepción, Chile. Docente con 5 años de experiencia en la Universidad Técnica Federico Santa María, especializado en las áreas de Robótica Industrial, CAD/CAM y Diseño Mecánico.

Profesores - mecanica.usm.cl

www.mecanica.usm.cl/10-profesores?start=15

Proyecto interno USM 25.15.11: ... Nuestros profesores y estudiantes de postgrado llevan a cabo además investigación fundamental y aplicada en los aspectos científicos de la disciplina, contribuyendo a la creación de nuevos conocimientos en los diversos campos de la espacialidad.

DTI • Universidad Técnica Federico Santa María

www.dti.usm.cl/servicios6

Para Alumnos, profesores y apoyos académicos vigentes. Wolfram Mathematica. Acceso a software Mathematica para uso académico en computadores de la Universidad. ... Copie los datos que se le presentan en la opción "Display Computer Information" y envíelos a mesa-deservicios.dti@usm.cl, con esto desde mesa de servicios le enviarán el archivo ...

Webmail Profesores/Funcionarios :: Log in - mail.usm.cl

<https://mail.usm.cl/horde>

- Procedimiento de obtención (o cambio) de clave de CUENTA USM - ¿ Dónde solicito la creación de una cuenta de email ? - Información sobre Correo Electrónico en la USM - Listas de Email (Mailling Lists) - Configuración Acceso IMAP/POP de Mozilla Thunderbird o MS Outlook 2003.

Ciencias USM | Quiénes Somos

ciencias.usm.cl/quienes-somos

El cuerpo docente cuenta con 12 profesores de planta adscritos al Departamento de Ciencias, de los cuales 5 profesores son de la especialidad de Matemáticas (cuatro de jornada completa de los cuales dos de ellos cuentan con cargo directivo, y uno media jornada adscrito),

3 profesores de la especialidad de Física (dos de jornada completa y ...

Profesores Jornada Completa - metalurgia.usm.cl

www.metalurgia.usm.cl/index.php/profesores/profesores-jornada-completa

Web del Departamento de Ingeniería Metalúrgica y de Materiales de la Universidad Técnica Federico Santa María

DTI • Universidad Técnica Federico Santa María

www.dti.usm.cl/servicios3

Debe conectarse a la red Wifi Profesores_USM; Identifíquese con su cuenta Institucional. Requisitos del servicio: Para alumnos, profesores o funcionarios vigentes. Conexión Wifi EDU-ROAM. Servicio de Conexión Wifi Eduroam dentro de la Universidad y en otras instituciones que cuentan con este servicio de conexión Wifi.

Profesores Postgrado - electronica.usm.cl

www.electronica.usm.cl/profesores-postgrado

Profesores del Programa (Actualizado Junio 2019) Claustro de Profesores del Programa(CDTP*) (*Cuerpo de Directores de Tesis del Programa) [Magister] Área de Especialización del Programa Computadores Fernando Auat C. Grado (año) Doctor (2009) Institución otorgante Universidad Nacional de San Juan, Argentina Línea de Investigación Robótica Móvil Gonzalo Carvajal B. Grado (año) PhD ...

6.4.2. Consulta: eclipse solar 2019

Eclipse solar del 2 de julio de 2019 - Wikipedia, la ...

https://es.wikipedia.org/wiki/Eclipse_solar_del_2_de_julio_de_2019

Eclipse solar del 2 de julio de 2019. El martes 2 de julio de 2019 se producirá un eclipse solar total que será visto en el océano Pacífico, Chile y al atardecer, en Argentina así como también en Uruguay.

Eclipse Solar Total 2 de Julio de 2019 - turismoastronomico.cl

<https://www.turismoastronomico.cl/eclipse-solar-total-2-de-julio-de-2019.html>

El 2 de julio de 2019 se producirá un eclipse solar total que será visible por una delgada franja del océano Pacífico y parte de América del Sur. La totalidad del eclipse comenzará al este de Nueva Zelanda en el Océano Pacífico para luego llegar a América del Sur. Los privilegiados esta vez serán los observadores de Chile y Argentina.

AURA - Chilean Eclipse 2019

<https://www.eclipsesolar2019.cl>

AURA - Chilean Eclipse 2019 La Asociación de Universidades para la Investigación en Astronomía (AURA) es un consorcio internacional de universidades e instituciones sin fines de lucro que operan instalaciones astronómicas de clase mundial .

Eclipse solar 2019: Esta será la duración en Chile del ...

<https://www.ahoranoticias.cl/noticias/tendencias/260560-eclipse-solar-2019-en-chile...>

Cada vez queda menos para el eclipse solar 2019 que podrá ser visto a la perfección en Chile, y las dudas sobre el fenómeno astronómico comienzan a surgir: desde por qué se ocasiona hasta cuánto durará.. Y si bien es un evento que pasa cada cierta cantidad de años en el país, los expertos han vuelto a hacer el llamado de lo peligroso que puede llegar a ver el eclipse sin estar ...

Eclipse Solar Total Argentina 2019 - Eclipses

www.eclipses.com.ar/eclipse_20190702.html

El eclipse solar del 2 de Julio de 2019. 1° JIPCCA. 1° Jornadas Internacionales de Promoción de la Cultura Científica en Astronomía. Más Información. Más información sobre el eclipse y sobre los horarios para su observación desde toda la Argentina en Miradas al Cielo y Sur Astronómico.

Eclipse solar 2019 en Chile - Astronomía en Chile y ...

<https://www.chileestuyo.cl/eclipse-solar-2019-en-chile-lugares-para-el-astroturismo>

Se realizarán observaciones de Eclipse desde el Cerro Patagual y el eclipse solar 2019 se verá en un 93,90 %. Muy próximo a la ciudad de Casablanca se encuentra el Observatorio Turístico Tokicura (4,6 km camino a Las Dichas). El eclipse en el Valle de Casablanca será parcial y se verá al 91,44 %.

Eclipse Solar Atacama 2019 - Atacama Eclipse Solar 2019

eclipsesolaratacama19.cl

El 2 de julio de 2019 se producirá un eclipse solar total que será visto en el océano Pacífico, Chile y al atardecer, en Argentina así como también en Uruguay. Un eclipse solar ocurre cuando la Luna pasa entre la Tierra y el Sol, oscureciendo total o parcialmente la imagen ...

TOUR ECLIPSE SOLAR 2019 La Serena, Higuera Chile 2 de ...

<https://sunnytravel.cl/eclipse-solar>

04-05-2019 · Eclipse Solar 2019 La Serena Tour en la higuera el 2 de julio de 2019 el mejor turismo ¡Únete a nosotros para la mejor experiencia de eclipse! RESERVA YA!

5/5

Autor: Sunnytravel

Total Solar Eclipse on July 2, 2019 - Time and DateTraducir esta página

<https://www.timeanddate.com/eclipse/solar/2019-july-2> 02-07-2019 · Total solar eclipse on Tuesday, July 2, 2019: Where and when is the Sun eclipse visible? Path map, animation, and local times.

Temporada de Eclipses

eclipseschile.gob.cl

Eclipse Atacama-Coquimbo 2 julio 2019. La Temporada de Eclipses se inicia con el fenómeno astronómico que tendrá su punto de oscuridad total desde la localidad de Domeyko en la región de Atacama, hasta el balneario de Guanaqueros en la región de Coquimbo, y que será visible, de manera parcial, en todo el territorio nacional.

Temporada de Eclipses

eclipseschile.gob.cl

Eclipse Atacama-Coquimbo 2 julio 2019. La Temporada de Eclipses se inicia con el fenómeno astronómico que tendrá su punto de oscuridad total desde la localidad de Domeyko en la región de Atacama, hasta el balneario de Guanaqueros en la región de Coquimbo, y que será visible, de manera parcial, en todo el territorio nacional.

Se confirma eclipse solar total en Chile el 2019 ...

[www.adnradio.cl/noticias/sociedad/se-confirma-eclipse-solar-total-en-chile-el-2019/...](http://www.adnradio.cl/noticias/sociedad/se-confirma-eclipse-solar-total-en-chile-el-2019/)

El evento natural que llamó la atención de los estadounidenses durante esta semana ya tiene fecha, e incluso hora, en nuestro país: se viene un eclipse solar total el próximo 2 de Julio de ...

Solar and Lunar Eclipses Worldwide – 2019Traducir esta página

<https://www.timeanddate.com/eclipse/2019>

LIVE Eclipse Stream. Watch our LIVE stream of the Total Solar Eclipse July 2, 2019. Go to LIVE Page. Transmisión en Vivo. Sigue nuestra transmisión en vivo del eclipse solar total 2 Julio, 2019 ...

Eclipse Solar 2019 ¿Cómo influirá en los signos del ...

<https://www.wemystic.com/es/eclipse-solar-2019>

Al hablar de Eclipse Solar 2019, nos referimos a los tres eventos astrológicos de este tipo, que se registrarán durante los meses de enero, julio y diciembre de 2019, y ...

Solar Eclipse 2019 Chile

eclipsesolar2019.com

Les invitamos a participar de una de las más extraordinarias experiencias de este siglo, el esperado eclipse solar 2019 a realizarse en el hermoso y místico Valle del Elqui.

Eclipse Solar 2019 - 24horas

<https://www.24horas.cl/eclipsesolar2019>

Científico chileno será uno de los 62 afortunados que verá el eclipse solar desde un avión . 62 personas podrán vivir el próximo eclipse solar del 2 de julio desde las alturas.

Solar eclipse of July 2, 2019 - WikipediaTraducir esta página

https://en.wikipedia.org/wiki/Solar_eclipse_of_July_2,_2019

A total solar eclipse will occur at the Moon's ascending node of the orbit on July 2, 2019 with a magnitude of 1.0459. A solar eclipse occurs when the Moon passes between Earth and the Sun, thereby totally or partly obscuring the image of the Sun for a viewer on Earth. A total solar eclipse occurs when the Moon's apparent diameter is larger than the Sun's, blocking all direct sunlight, turning ...

Conoce el recorrido del Eclipse Total de Sol que se verá ...

www.diarioeldia.cl/region/conoce-recorrido-eclipse-total-sol-que-se-vera-en-serena-en-2019

“Nos queda prepararnos para lo que viene en 2019, ya tenemos la capacidad hotelera casi al 100 % reservada y nos queda preparar a nuestros estudiantes para que ellos sean los guías turísticos, para que le cuenten al turista qué significa este eclipse solar.

Eclipse solar 2020 - Araucanía | Eclipses Chile 2019 - 2020 v0

<https://www.conicyt.cl/eclipseschile1/eclipse-solar-2020-araucania>

El lunes 14 de diciembre de 2020, un eclipse total de Sol cruzará la Región de La Araucanía. En el resto del territorio nacional podrá observarse de manera parcial. El sur de Chile es la cuna de la cultura Mapuche, donde la observación de los astros y su influencia en la vida cotidiana, dieron origen a ...

Mapa del Eclipse - AURA - Chilean Eclipse 2019

<https://www.eclipsesolar2019.cl/eclipse/mapa-del-eclipse>

Credito: Observatorio Nacional Solar de EEUU, AURA y la Fundación Nacional de Ciencias de EE.UU. Usando la franja del Eclipse de NASA generado por Fred Espenak, Bessleian Elements para el Eclipse de Julio, 2019.

6.4.3. Consulta: radio carolina

Radio Carolina

www.carolina.cl

A pesar de ganarle a Tailandia por 2-0, le faltó un gol más para clasificar a octavos de final.

Radio Carolina online - Señal en vivo - 99.3 MHz FM ...

<https://onlineradiobox.com/cl/carolina>

Radio Carolina 99.3 MHz FM, Santiago de Chile, Chile - escuchar radio online gratis en On-lineRadioBox.com. Esta página web utiliza cookies. Al continuar utilizando esta página web, estás de acuerdo con nuestras políticas respecto al uso de cookies.

RADIO CAROLINA - ON LINE - SEÑAL EN VIVO - http://www ...

www.talcaguia.cl/radio_online.php?id=19

radio carolina en vivo señal online gratis chile en directo - on line señal en vivo en directo
<http://www.carolina.cl/>

Radio Carolina, 99.3 FM, Santiago, Chile | Free Internet ...

<https://tunein.com/radio/Radio-Carolina-993-s25096>

Radio Carolina - Hiperkinética y de permanente buen humor, Carolina trae lo mejor de los sonidos de moda en una selección que incluye música nacional, Electronica, Pop y Reggaetón.

Radio Carolina - Wikipedia, la enciclopedia libre

https://es.wikipedia.org/wiki/Radio_Carolina

A comienzos de 2009 Radio Carolina abandona la frecuencia 100.3 de San Antonio, frecuencia ocupada hasta 2017 por Radio Beethoven, provocando molestia entre sus auditores de esa zona. Desde la radio no dieron ningún tipo de explicaciones sobre su salida.

Eslogan: La Más Prendida, Yo Escucho Carolina

Radios Online Chile, emisoras chilenas, radios en vivo

<https://emisora.cl> ¿Buscas una radio online de Chile? En nuestra web encontrarás tu radio en vivo favorita, ¡con toda la música online a un solo clic! ¡Fácil, rápido y ameno!

Radio Carolina - Señal Online - El Telón - Radio y ...Traducir esta página

www.eltelon.com/radios/carolina

Disfruta Radio Carolina online las 24hrs del día. Descubre cientos de radios en vivo en nuestro sitio.

Mi Pagina, Portal de Sitios Chilenos

mipagina.cl

Miles de sitios chilenos útiles , clasificados en categorías. Todos los Diarios y radios de Chile. .Msn Messenger via web. Escuchar todas las radios online de Chile . Traductor. Medidor de conexion

Mi Pagina, Portal de Sitios Chilenos

mipagina.cl

Miles de sitios chilenos útiles , clasificados en categorías. Todos los Diarios y radios de Chile. .Msn Messenger via web. Escuchar todas las radios online de Chile . Traductor. Medidor de conexion

Radio Carolina online - Escuchar Radio on line

<https://radiochilena.net/radio-carolina>

Radio Carolina. Radio Carolina es una de las estaciones preferidas por la juventud chilena, ya que incorpora música urbana actual tales como reggaetón, hip-hop, EDM, dance y hasta pop.

Radio Carolina - guiapenquista.cl

www.guiapenquista.cl/radios-chilenas/radio-carolina.php

Radio Carolina online señal en vivo frecuencia 99.3 Fm por internet para Chile y el Mundo

Radio Karolina - Lagana i fina

<https://karolina.rs>

Radio Karolina je prisutan u srpskom medijskom prostoru od 2013. godine. Najpre je bila pokrivena samo teritorija Beograda, da bi u međuvremenu i Novi Sad i Niš dobili svoje frekvencije. Radio je koncipiran tako da se slušaocima, ali i tržištu uopšte, ponudi program koji je jedinstven u Srbiji.

Radio CarolineTraducir esta página

www.radiocaroline.co.uk

Radio Caroline, the former offshore radio station, now sending multiple music streams on the internet, also local full time DAB and regional AM on special weekends

Radio Carolina - Señal Online - eltelon.com

www.eltelon.com/radios/carolina?in=chile

Disfruta Radio Carolina online las 24hrs del día. Descubre cientos de radios en vivo en nuestro sitio.

RADIO CAROLINA - Home | Facebook Traducir esta página

<https://www.facebook.com/RADIOCAROLINA>

RADIO CAROLINA, Ñuñoa. 1,158,790 likes · 79,641 talking about this. Radio Carolina! 🎧 La Más Prendida 🎧

Uživo Radio Karolina Beograd, Karolina radio streaming ...

www.uzivoradio.com/karolina-beograd.html

Radio Karolina lagana i fina iz Beograda uzivo. Slušajte Karolina radio uživo, internet radio stream. We use cookies to ensure that we give you the best experience. If you continue using this website, we'll assume that you are OK about that.

Radio Carolina - En vivo en línea Radio

<https://www.liveonlineradio.net/es/chile/radio-carolina.htm>

Radio Carolina basado en La Serena, Chile es uno de la estación de la música popular. Estación de Radio Carolina música en streaming y Programms tanto en el aire y en línea. Originalmente es una música española, Adulto Contemporáneo, Top canal de radio 40 / ...

Daddy Yankee - Dura (Video Oficial) - YouTube

6.4.4. Consulta: red transantiago

Red Metropolitana de Movilidad

<https://www.red.cl>

Con nuevos servicios y extensiones, Ministerio de Transportes estrena malla de recorridos que beneficiará a nueve comunas de Santiago. Los cambios, que debutan el domingo, contemplan tres nuevos recorridos, dos ampliaciones y once ajustes que darán más conectividad a Cerrillos, Estación Central, Santiago, Pudahuel, Quilicura, Vitacura, Ñuñoa, La Reina y Peñalolén.

Mapas de recorridos | Red Metropolitana de Movilidad

<https://www.red.cl/mapas-y-recorridos/mapas-de-recorridos>

A continuación encontrarás el mapa de las 32 comunas de la Región Metropolitana, más San Bernardo y Puente Alto, junto con las rutas habituales de todos los recorridos de buses de las empresas concesionarias que operan en estas comunas.

Las características de Red que lo hacen distinto al ...

<https://www.fayerwayer.com/2019/03/red-transantiago-caracteristicas>

12-03-2019 · Red ya comienza su implementación y lo que queda de acá en más es descubrir exactamente qué diferencias existirán con el famoso y tan vilipendiado Transantiago. Según

lo que han prometido las autoridades, este sistema se fundamenta en contratos más cortos y flexibles, con empresas de buses ...

Dale dale RED: Así es el nuevo sistema de transporte que ...

<https://www.fayerwayer.com/2019/03/red-transantiago-transporte>

12-03-2019 · RED, ese es el nombre del nuevo sistema de transporte que reemplazará al Transantiago y que este martes 12 de marzo presentó el presidente Sebastián Piñera y la ministra Gloria Hutt. ???Hola Bienvenidos. Te invitamos a conocer @Red_Movilidad el ...

De Transantiago a Red: ¿Solución real o mera estrategia ...

<https://radio.uchile.cl/2019/03/12/de-transantiago-a-red-solucion-real-o-mera...>

El presidente Sebastián Piñera anunció, este martes, que el Transantiago cambiará de nombre para ser rebautizado como “Red Metropolitana de Movilidad”. Según comentó el mandatario, la nueva imagen se debe a que será una “red” que conectará los buses, el Metro, Metrotren y bicicletas ...

Autor: Rodrigo Fuentes

Experto prevé alza de tarifas de Red (ex Transantiago ...

<https://www.cooperativa.cl/noticias/pais/transportes/transantiago/experto-preve-alza...>

Nuevos recorridos del Transantiago beneficiarán a 700 mil personas Gobierno dio el primer paso formal en cambio de Transantiago a Red Santiago tendrá el primer corredor 100 % eléctrico de ...

Nuevo sistema de transporte público: las diferencias entre ...

<https://www.24horas.cl/nacional/nuevo-sistema-de-transporte-publico-las-diferencias...>

12-03-2019 · Este martes, el Presidente Sebastián Piñera junto a la ministra de Transportes, Gloria Hutt, presentó Red”, el nuevo sistema de transporte público que reemplaza al Transantiago y anunció la ...

Ubicación: Bellavista 0990, Providencia, Santiago, 832 0000

Red Metropolitana de Movilidad - Wikipedia, la ...

<https://es.wikipedia.org/wiki/Transantiago> Red Metropolitana de Movilidad (conocido simplemente como Red; hasta marzo de 2019 denominado Transantiago) [8] es un sistema de transporte público urbano que opera en el área metropolitana de la ciudad de Santiago, capital de Chile. Integra servicios de buses con el Metro de Santiago y el servicio de tren suburbano Metrotren Nos. Su ...

Medios de transporte: Buses, Metro, Tren suburbano

Ubicación: Santiago, Chile

Tipo: Transporte público Longitud: Buses: 2834 km[4], Metro: 140 km[5], Metro-tren: 20 km[6] Red (Transantiago) | Interior Bus de 2 pisos ...Traducir esta página https://www.youtube.com/watch?v=3_JHyuaaFPw

6:22 YOUTUBE

27-03-2019 · #RedMejorTransporte #Wrightbus #BusesVule Grabación de exterior e interior

de un bus a

prueba del transporte público de Santiago de Chile, unidad Wrightbus Streetdeck 10.6 Euro 6. Este bus está ...

Autor: Ariel YOUTUBE

Visualizaciones: 8,1K

Primer día de circulación Alstom Aptis - Interior y ...

«Red»: El sistema de transportes que reemplazará al ...

<https://www.rockandpop.cl/2019/03/red-el-sistema-de-transportes-que-reemplazara-al...>
12-03-2019 · 12 marzo, 2019 «Red»: El sistema de transportes que reemplazará al Transantiago El plan sumará nuevos buses, nuevas líneas de Metro, inclusión de bicicletas y la transformación en la forma de pago de la tarjeta Bip!.

De Transantiago a Red – Noticias UAI

<https://noticias.uai.cl/de-transantiago-a-red>

El presidente Sebastián Piñera anunció que el Transantiago cambiará de nombre a “Red Metropolitana de Movilidad”, y comentó que la nueva imagen se debe a que será una “red” que conectará los buses, el metro, metrotren y bicicletas, y que el cambio llegará con la puesta en marcha de 200 buses eléctricos y 490 ecológicos, lo que significaría un mejor estándar de servicio en ...

Red: el nuevo nombre que tendrá el Transantiago - t13.cl

<https://www.t13.cl/noticia/nacional/red-nuevo-nombre-tendra-transantiago>

Red”: el nuevo nombre que tendrá el Transantiago. Presidente Piñera realizará este martes una ceremonia para presentar el “cambio de identidad” que tendrá ...

Cambio de Transantiago a Red renovará el 70 % de los buses

<https://www.rockandpop.cl/2019/03/transantiago-red-renovara-buses>

13 marzo, 2019 Cambio de Transantiago a Red renovará el 70 % de los buses La modificación al sistema de transportes de la capital, que pasará después de más de una década como Transantiago a Red, cambiará los buses.

Así funcionará Red, el nuevo sistema de transporte capitalino

<https://www.lacuarta.com/cronica/noticia/red-transantiago-capital/346111>

13-03-2019 · “Red es mucho más que una marca, es mucho más que un servicio, es realmente un compromiso de unir a los chilenos a través de un nuevo sistema de transporte público, que les mejore su calidad de vida y les permita una vida más plena y más feliz”, ...

Transantiago tendrá un nuevo nombre: se llamará Red ...

<https://www.biobiochile.cl/noticias/nacional/region-metropolitana/2019/03/12/aseguran...>

El Gobierno del presidente Sebastián Piñera ha prometido en reiteradas ocasiones una transformación completa al Transantiago, sistema de transportes que prometía mejorar la forma de ...

Las reacciones que ha dejado Red”, el reemplazo del ...

<https://www.fmdos.cl/noticias/reacciones-red-sistema-reemplaza-transantiago>

Este martes el presidente Sebastián Piñera presentó el nuevo sistema de transporte «Red» (Red Metropolitana de Movilidad), que viene a reemplazar al Transantiago tras 11 años de su funcionamiento.. Fueron 550 millones de pesos invertidos en la asesoría comunicacional que buscó darle un nuevo rostro al transporte público de la capital, según informa «La Tercera».

Nueva flota de buses y paraderos inteligentes: Red, el ...

https://www.cnnchile.com/pais/red-sistema-transporte-publico-transantiago_20190312

Nueva flota de buses y paraderos inteligentes: Red, el sistema de transporte que reemplazará al Transantiago. ... El nuevo concepto Red denominará diversos cambios en el sistema de transportes.

Red: El fin del Transantiago - Libertad y Desarrollo

<https://lyd.org/centro-de-prensa/noticias/2019/03/red-el-fin-del-transantiago>

Red: El fin del Transantiago. 13 de marzo de 2019 Facebook. Twitter. Linkedin. El Presidente de la República, Sebastián Piñera, junto con la Ministra de Transportes y Telecomunicaciones, dieron a conocer el cambio de identidad que va a tener el transporte público de Santiago, que pasará a llamarse Red Metropolitana de Transporte, o ...

En qué consistirá Red”: El nuevo sistema de transporte ...

<https://www.ahoranoticias.cl/noticias/nacional/253643-red-cambio-transantiago-modifica...>

Nacional En qué consistirá Red”: El nuevo sistema de transporte capitalino que reemplazará al Transantiago Entre las novedades destaca el cambio de colores y estándares para operadores.

6.4.5. Consulta: comprar caliente camas

Caliente Camas en Mercado Libre Chile

<https://hogar.mercadolibre.cl/dormitorio/ropa-cama/caliente-camas>

Encuentra Caliente Camas en Mercado Libre Chile. Descubre la mejor forma de comprar online.

Ofertas en Todos las Calientacamás | Abcdin.cl - Tienda Online

<https://www.abcdin.cl/tienda/es/abcdin/calientacamás>

Calientacamás en ofertas a precios imperdibles, calienta tu cama en Abcdin.cl | La felicidad cuesta menos Aquí! Cerrar Comparar Agrega otro Producto ... Comprar Ver detalle Comparar Comparar Calientacama Scaldasonno Matrimonial 150 x 160 cms Internet \$52.990 Normal \$54.990

Calientacamás | Ripley.com

<https://simple.ripley.cl/electro/climatizacion/calientacamás>

Comprar en Ripley.com es 100 % seguro. Comercial ECCSA S.A., Rut N° 83.382.700-6, domiciliada en Avenida Cerro Colorado 5240, Torre del Parque 1 Piso 11, comuna Las Condes, Santiago de Chile. Precios publicados aplican exclusivamente a Internet y Venta Telefónica, estos pueden llegar a ser distintos en tiendas.

Comprar Calientacamás | Ofertas, precios y catálogos

<https://www.tiendeo.cl/ofertas/calientacamás>

Comprar calientacamás - Ofertas y promociones Favorito . Ofertas. Ofertas (5) Catálogos (5) Cupones (0) Catálogos de Calientacamás y de Hogar y muebles en tu ciudad. Ha habido un problema inténtelo más tarde. ¡Añade París a tus Favoritos! Catálogo París En Tienda Válido 11 días.

Calientacamás | Sodimac.com

<https://www.sodimac.cl/sodimac-cl/category/scat299483/Calientacamás>

Encuentra todo lo que buscas en Calientacamás al mejor precio. Visita Sodimac.com y conoce nuestra gran variedad.

Calientacamás - Elige comodidad y calidez al dormir | París.cl

<https://www.paris.cl/dormitorio/ropa-cama/calientacamás>

Olvidate del frío con estos prácticos calienta camas y disfruta de un sueño confortable sin preocuparte de las bajas temperaturas. Búscalos en París.

Scaldasonno Ropa De Cama Calienta Camas - Todo para ...

<https://hogar.mercadolibre.cl/dormitorio/scaldasonno-ropa-de-cama-calienta-camas>

Encuentra Scaldasonno Ropa De Cama Calienta Camas - Dormitorio en Mercado Libre Chile. Descubre la mejor forma de comprar online.

Amazon.es: Calientacamás y cubrecolchones eléctricos ...

<https://www.amazon.es/Calientacamás-Cubrecolchones-Eléctricos-Masaje-Relajacion/b?ie=...>

Compras en línea de Calientacamás y cubrecolchones eléctricos de una gran selección en la tienda Salud y cuidado personal. Compras en línea de Calientacamás y cubrecolchones eléctricos de una gran selección en la tienda Salud y cuidado personal. Saltar al contenido principal.

Calientacamás Kioto 2 Plazas Digital KTC-1006 | ABCDIN

<https://www.abcdin.cl/tienda/es/abcdin/calientacamás-kioto-2-plazas-digital-ktc-1006...>

Calientacamás Kioto 2 Plazas Digital KTC-1006; Calientacamás Kioto 2 Plazas Digital KTC-1006. \$16.990 Normal \$ 16.990 Internet Medios de pago disponibles ... Comprar Producto Agotado Despacho de Productos Infórmate acerca de cuánto es el costo de despacho de este producto. El tamaño para este cálculo es: Pequeño.

Comprar calientacamás en San Joaquín - Ofertas ...

<https://www.tiendeo.cl/san-joaquin/ofertas/calientacamás>

Comprar calientacamás en San Joaquín - Ofertas, promociones y cupones Seguir . Ver oferta

. 18 días más. Especial belleza . 1. Bolsa de agua caliente. Ver oferta . 4 días más. Especial día del padre . 1. Calientapiés. Ver oferta . 18 días más. Aprovecha tus puntos y ahorra . 1. Manta eléctrica. Ver oferta .

Calienta Camas al mejor precio en Mercado Libre Uruguay

<https://listado.mercadolibre.com.uy/hogar-muebles/dormitorio/ropa-cama/calienta-camas>

Encontrá los mejores precios en Calienta Camas en Mercado Libre Uruguay.¡Más de 602 ofertas increíbles te están esperando! Mercado Libre Uruguay - Donde comprar y vender de todo Bienvenido

Calienta camas | Hites

<https://www.hites.com/electro-hogar/climatizacion/calientacamass>

Compra Calienta camas en nuestras tiendas Hites. En Hites.com podrás equipar tu casa en electrohogar con los mejores Refrigeradores, Lavadoras, Secadoras, Cocina, Cuidado Personal, Electrodomésticos y Climatización.

Comprar calentacamass y cubrecolchones baratos - Orbegozo ...

<https://www.electrocosto.com/calientacamass>

¿Qué calentacamass eléctrico comprar? Consejos y Recomendaciones. Las mejores calentacamass eléctricos. Las mejores ofertas en calentacamass eléctricos las vas a encontrar en Electrocosto. En invierno si quieres proteger tu cama del frío y la humedad el cubrecolchones o calentacamass ...

Comprar Calientacamass en Curicó - Ofertas y tiendas - Ofertia

<https://www.ofertia.cl/curico/ofertas/calientacamass>

¡Encuéntralas en Ofertia! Todas las tiendas en Curicó con catálogo de Calientacamass. Accede sencillamente a catálogos, ofertas y folletos de las tiendas a tu alrededor. Producto, marca o nombre de la tienda. Ciudad. Buscar. ... Tiendas donde comprar Calientacamass en Curicó ...

Comprar un Calientacamass - Manta eléctrica

tumantaelectrica.com/calientacamass-mas-vendidos

Te presento una breve comparativa de los calentacamass eléctricos calefactables mejor valorados por los usuarios en Amazon. Para no pasar frío en la cama te recomiendo estos calentacamass: tienen la mejor relación calidad-precio. En mi blog, te recomiendo mantas eléctricas y almohadillas cervicales.

Qué son los calentacamass y consejos para comprarlos ...

<https://www.bricolemar.com/blog/que-son-los-calientacamass>

Ya os he dicho qué es un calentacamass y cómo son fabricados para sacar el máximo rendimiento a las resistencias eléctricas para producir calor y mantener la cama muy agradable durante las estaciones frías como otoño e invierno. Ahora bien, os daré una serie de consejos a tener en cuenta antes de comprar un calentacamass

Mejor Calientacamass Eléctrico?junio 2019 ?-Comparativa

<https://consejosparaahorrar.net/hogar-y-cocina/manta-electrica/mejores-calientacamamas...>
A continuación puede leer nuestra guía de compra para conocer sus características, tipos y saber comprar las mejores calientacamamas térmicas en base a criterios de calidad, precio, ofertas e información correcta para el consumidor, pero de todas formas, si no quiere usar ese tiempo, le dejamos la lista de las mantas eléctricas para camas mas vendidas del año en Amazon!.

Comprar Calientacamamas | Ofertas y promociones

<https://www.tiendeo.com/ofertas/calientacamamas>

Los calientacamamas son una serie de productos ideados para calentar la cama y proporcionar un mayor confort. Antiguamente el calentador de cama era un recipiente de cobre o latón en el que se introducían piedras o brasas calientes. Gracias a su diseño, que contaba con un largo mango, podían moverse por el interior de la cama para distribuir bien el calor.

CALIENTACAMAS - Siegen

<https://siegen.cl/15-calientacamamas>

Te ayudamos con tu compra +569 4433 1951 +562 2512 0184. 0 producto Productos vacío.
Ningún producto

calienta camas - Comparar precios y opiniones - Comprar

<https://www.bestalia.com/comprar/calienta-camas>

calienta camas - Comparar precios y opiniones para Sanitas SWB 50 - Manta térmica para la cama [importado de Alemania], Cubrecolchon Electrico Fab En Europa 2 Zonas Calienta Camas Tipo Manta Electrica, Bosch - Calientacamamas Pfb2030 (150 x 80 cm), Calienta camas astan ah51020 (160x140cm) doble, Calientacamamas Electrico Çalentar Cama "150x80 Cm Calor Manta Almohadilla, .

6.4.6. Consulta: 13c vivo online

En vivo Canal 13 | 13.cl

<https://www.13.cl/en-vivo>

Sitio web oficial de Canal 13, primer canal de transmisión chileno en transformarse en una casa editorial que traspasa sus contenidos a todas sus plataformas como Internet, radio, cable y ...

13C | 13.cl

<https://www.13.cl/c>

Inés Matte Urrejola 0848, Santiago, Chile - Fono (562) 2 251 4000 © Todos los derechos reservados. 13.cl

**Inicio • Días De Música • Crónicas De Chile • Documentales • Corte Nacional • Ojo En Tinta
Canal 13 Cable En Vivo | TV Online Chile**

tv-online.cl/canal-13-cable-en-vivo.php

Canal 13 Cable transmite de forma ininterrumpida las 24 horas del día y gracias a tv-online.cl podrás ver su señal en directo desde tu teléfono, computador o Tablet. No esperes más y comienza a disfrutar de todo el contenido de Canal 13C hoy mismo y ve la cultura con otros ojos!

Canal 13 En Vivo | TV Online Chile

tv-online.cl

Canal 13 En Vivo, TV Online Chile: Sintoniza todas las señales abiertas de Televisión Chilena en Internet

Canal 13 C Chile en vivo, Online Teleame Directos TV

directostv.teleame.com/canal13cchileenvivo

Canal 13 C Chile en vivo, Online, en directo. 13C (anteriormente Señal 3 y Canal 13 Cable) es un canal de televisión por cable chileno, de propiedad y operado por Canal 13 S.A.

Autor: Teleame

Deportes 13 - En Vivo

envivo.13.cl Can't find variable: URL. Seguir a @t13 # MundialxEl13

Señal Online - Chilevisión

www.chilevision.cl/senal-online

Disfruta del contenido de Chilevisión en nuestra señal online. Somos el canal oficial de la Selección Chilena de Fútbol y #VamosContigo

Canal13C - YouTube

<https://www.youtube.com/channel/UCLgzQavukZG7Q4Chi7WjsQg>

Somos 13C, un canal de televisión de pago que explora la cultura, los viajes, la vida sana, los personajes, las tendencias y la actualidad. Estamos en Chile,...

Canal 13 en vivo - (El trece) Mirá todos los canales de TV ...

www.televisionparatodos.tv/canal-13-en-vivo-online

Canal 13 en vivo, actualmente conocida por su nombre comercial El Trece, es una cadena de televisión argentina que transmite desde la Ciudad Autónoma de Buenos Aires, siendo la emisora con mayor cobertura nacional. Es la encargada de proporcionar contenidos de programación a diferentes canales de televisión en el interior del país, a través de su señal satelital.

Canal 13 - YouTube Traducir esta página

<https://www.youtube.com/user/canal13>

Lo mejor de Canal 13 está en YouTube ¡Suscríbete y disfruta semana a semana nuestros mejores momentos!

TV online, Canales de Televisión de Chile, Señal en vivo

www.refnet.cl/tvonline.html

TV online, Señal en vivo Canales de Televisión de Chile, Windows Media, Flash Player

CANAL 13 | ON LINE SEÑAL EN VIVO EN DIRECTO GRATIS

www.talcaguia.cl/tv_online.php?id=8

canal 13 | on line señal en vivo gratis. canal 13 | on line señal en vivo gratis ... además, posee desde 1995 una segunda señal en el cable, llamada 13c (antes señal 3), con programación cultural, y tiene a su cargo teleduc, creado en 1977, que es el centro de educación a distancia de la universidad católica.

En Vivo - eltrece tv.com.ar

<https://www.eltrecetv.com.ar/vivo>

Mirá la programación en vivo de El Trece.

TVN En Vivo - Televisión Nacional de Chile

<https://www.tvn.cl/envivo/?service=mobile&device=mobile-bot>

TVN En Vivo, Sitio web de Televisión Nacional de Chile, TVN. Somos el canal de Chile y nuestra programación ofrece programas de noticias de Chile y el mundo, reportajes, debate, servicio, telenovelas, series, películas, cultura, espectáculos y deporte.

[EN VIVO] Sigue la transmisión de Tele13 Radio | Tele 13

<https://www.t13.cl/radio/en-vivo/videos/en-vivo-senal-video-de-tele13-radio>

EN VIVO - Sigue la transmisión de Tele13 Radio. Página 13. Cavallo y Martínez por "Dolor y Gloria" "La Desaparición" Club de Fin de Semana. Comidas "de la mitad del mundo" la peor desgracia de la historia discográfica de EE.UU. Protagonistas.

Ver Canal 13 Cable en vivo - Canal 13 Cable online

<https://vertelevisionenvivo.blogspot.com/2013/03/canal-13-cable.html>

Canal 13 Cable de Chile en vivo por internet. Señal Internet del Canal de la Pontificia Universidad Católica de Chile, uno de los más populares del país de propiedad mixta (privado y universitario), regido por una sociedad por acciones, y que transmite por la frecuencia 13 de Santiago.

T13 Chile en vivo, Online Teleame Directos TV

directostv.teleame.com/t13envivo

Ver T13 Chile en vivo, Online, en directo, live, on-line, en línea, en internet, live streaming. Teletrece Chile. T13 Móvil.

ON LINE SEÑAL EN VIVO EN DIRECTO GRATIS - talcaguia.cl

www.talcaguia.cl/tv_online.php?id=181

Eleodoro Rodríguez Matte, director ejecutivo en 1995, decide internacionalizar la señal Canal 13, la estación más vista a nivel nacional. Para ello inicia la creación de dos frecuencias paralelas a la señal abierta: señal internacional, vía satélite; y la señal 3, vía cable.

Televisión Nacional de Chile

<https://www.tvn.cl>

Sitio web de Televisión Nacional de Chile, TVN. Somos el canal de Chile y nuestra programación ofrece programas de noticias de Chile y el mundo, reportajes, debate, servicio, teleno-

velas, series, películas, cultura, espectáculos y deporte.

Canal 13 Cable en Vivo | La Mejor | TV de Chile

lamejor.cl/canal13cable.html

Canal 13 Cable en Vivo. Dejar un comentario del programa que estas viendo Los comentarios publicados son de responsabilidad de quienes los emiten, no representan el pensamiento de lamejor.cl. No obstante, nos reservamos el derecho de administrar las listas de comentarios cuando éstos presenten ofensas personales.

REFERENCIAS BIBLIOGRÁFICAS

- [Albornoz et al., 2017] Albornoz, M. C., Berón, M., y Montejano, G. A. (2017). Interfaz gráfica de usuario: el usuario como protagonista del diseño. En *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*.
- [Alexa, 2019] Alexa (2019). Top Sites in Chile. <https://www.alexa.com/topsites/countries/CL>. [Online; Accedido el 17-Febrero-2019].
- [Broder, 2002] Broder, A. (2002). A taxonomy of web search. En *ACM Sigir forum*, volumen 36, pp. 3–10. ACM.
- [educativa Partesdel.com, 2018] educativa Partesdel.com, R. (2018). Partes del URL. https://www.partesdel.com/partes_de_la_direccion_web.html. [Online; Accedido el 22-Mayo-2019].
- [Garver, 1998] Garver, S. M. (1998). What is a Makefile? <http://www.sis.pitt.edu/mbsclass/tutorial/advanced/makefile/whatis.htm>. [Online; Accedido el 31-Mayo-2019].
- [Goyvaerts, 2017] Goyvaerts, J. (2017). Regular Expressions Tutorial. <https://www.regular-expressions.info/tutorial.html>. [Online; Accedido el 21-Mayo-2019].
- [Heritrix, 2013] Heritrix (2013). Heritrix Wiki. <https://github.com/internetarchive/heritrix3/wiki>. [Online; Accedido el 29-Abril-2019].
- [Hofmann, 2001] Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine learning*, 42(1-2):177–196.
- [Hofmann, 2004] Hofmann, T. (2004). Latent semantic models for collaborative filtering. *ACM Transactions on Information Systems (TOIS)*, 22(1):89–115.
- [Huang et al., 2003] Huang, C.-K., Chien, L.-F., y Oyang, Y.-J. (2003). Relevant term suggestion in interactive web search based on contextual information in query session logs. *Journal of the American Society for Information Science and Technology*, 54(7):638–649.
- [Huang et al., 2004] Huang, Z., Chen, H., y Zeng, D. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems (TOIS)*, 22(1):116–142.
- [Huang et al., 2007] Huang, Z., Zeng, D. D., y Chen, H. (2007). Analyzing consumer-product graphs: Empirical findings and applications in recommender systems. *Management science*, 53(7):1146–1164.
- [ISO, 2009] ISO, I. (2009). 28500: 2009 information and documentation-warcs file format. *International Organization for Standardization*.

- [Kowalska, 2016] Kowalska, M. (2016). Vertical Search vs Horizontal Search. <http://www.advertangle.com/vertical-search-vs-horizontal-search/>. [Online; Accedido el 14-Diciembre-2018].
- [Liu, 2007] Liu, B. (2007). *Web data mining: exploring hyperlinks, contents, and usage data*. Springer Science & Business Media.
- [Mendoza et al., 2009] Mendoza, M., Guerrero, H., y Farias, J. (2009). Inquirol: a new search engine in Chile. WWW'09.
- [Navia, 2015] Navia, W. (2015). Etiquetas para estructurar texto HTML5. <https://es.slideshare.net/WendyNaviaChambi/etiquetas-para-estructurar-texto-html5-2-parte/3>. [Online; Accedido el 29-Mayo-2019].
- [Netcraft, 2019] Netcraft (2019). June 2019 Web Server Survey. <https://news.netcraft.com/>. [Online; Accedido el 17-Junio-2019].
- [ScrapingHub, 2008] ScrapingHub (2008). Scrapy. <https://scrapy.org/>. [Online; Accedido el 29-Abril-2019].
- [Venners, 1996] Venners, B. (1996). The lean, mean, virtual machine. <https://www.javaworld.com/article/2077184/the-lean--mean--virtual-machine.html>. [Online; Accedido el 30-Mayo-2019].