

**ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES
EN PÁGINAS WEB**

EMIRALDO ANDRÉS LOZANO SÁNCHEZ

**UNIVERSIDAD DE LOS LLANOS
FACULTAD DE CIENCIAS BÁSICAS E INGENIERÍA
ESCUELA DE INGENIERÍAS
PROGRAMA DE INGENIERÍA DE SISTEMAS
VILLAVICENCIO, COLOMBIA
2017**

**ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES
EN PÁGINAS WEB**

EMIRALDO ANDRÉS LOZANO SÁNCHEZ

Trabajo de grado presentado como requisito parcial para optar al título de Ingeniero de sistemas

Director:

M.Sc. Roger Calderón Moreno

Codirector:

M.Sc. Javier Mauricio Enciso Moreno

**UNIVERSIDAD DE LOS LLANOS
FACULTAD DE CIENCIAS BÁSICAS E INGENIERÍA
ESCUELA DE INGENIERÍA
PROGRAMA DE INGENIERÍA DE SISTEMAS
VILLAVICENCIO, COLOMBIA
2017**

NOTAS DE ACEPTACIÓN:

Roger Calderón Moreno
Director

Javier Mauricio Enciso Moreno
Co-Director

Juan Fajardo Barrero
Jurado

Fecha

AGRADECIMIENTOS

Agradezco a todas las personas que hicieron posible la realización de este proyecto.

A mi familia por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación, por brindarme la oportunidad de ejecutar mis estudios de pregrado, pero más que nada, por su amor.

A mis compañeros de estudio, a mis maestros y amigos, quienes sin su ayuda nunca hubiera podido ejecutar con éxito este proyecto. Al ingeniero Javier Enciso, por la oportunidad, apoyo y supervisión durante todo el trabajo de pasantía. Así mismo, agradezco al ingeniero Roger Calderón Moreno, por su apoyo y supervisión durante el desarrollo de la pasantía en su rol como director de proyecto.

Agradezco a la empresa MOBILE CORP por la oportunidad prestada y el apoyo brindado a lo largo de la pasantía.

**TABLA DE CONTENIDO**

RESUMEN.....	4
JUSTIFICACIÓN.....	4
OBJETIVOS.....	5
OBJETIVO GENERAL	5
MARCO DE REFERENCIA.....	5
REFERENCIA TEÓRICA.....	6
TRABAJOS SIMILARES	7
DESARROLLO.....	8
RECOLECCIÓN DE INFORMACIÓN	8
<i>Búsqueda de plataformas con objetivos similares.....</i>	<i>8</i>
<i>Análisis de los algoritmos de búsqueda y recuperación utilizados en las plataformas encontradas.....</i>	<i>8</i>
DEFINICIÓN Y CLASIFICACIÓN DE ESTRATEGIAS DE BÚSQUEDA.....	8
<i>Definición de estrategias de búsqueda.</i>	<i>8</i>
<i>Clasificación por complejidad de dichas estrategias.</i>	<i>9</i>
SELECCIÓN DE TECNOLOGÍAS Y ESTRATEGIAS A IMPLEMENTAR.....	9
DISEÑO Y DESARROLLO DE ESTRATEGIAS	10
<i>Diseño</i>	<i>10</i>
<i>Orden de ejecución de estrategias.....</i>	<i>14</i>
<i>Desarrollo.....</i>	<i>14</i>
Implementación de los algoritmos	14
Dificultades encontradas en la implementación de estrategias.....	14
Desarrollo de ejecución de los algoritmos en Segundo plano.	15
Conexión con software empresarial.....	15
Comparación de rendimiento de las estrategias implementadas.....	15
Definición de exactitud de la búsqueda que realiza el sistema.	17
RESULTADOS OBTENIDOS.....	18
CONCLUSIONES.....	18
RECOMENDACIONES	19
REFERENCIAS.....	19
APÉNDICE A. ALGORITMO PLANTEADO.....	21
APÉNDICE B. CONEXIÓN CON SOFTWARE EMPRESARIAL.....	25
APÉNDICE C. EVIDENCIA RESULTADOS.....	34



APÉNDICE D. REGISTRO DE SOFTWARE.....	49
APÉNDICE E. PRESENTACIÓN EN VILLADEVS	50
APÉNDICE F. MENCIÓN A BYE BYE 404.....	52

LISTA DE TABLAS

Tabla 1. Comparación estrategias. Fuente: Autor.....	16
Tabla 2. Efectividad búsqueda. Fuente: Autor.....	17
Tabla 3. Evidencia resultados obtenidos con algoritmo. Fuente: Autor.	34

LISTA DE ILUSTRACIONES

Ilustración 1, Arquitectura de Solución. Fuente: Propia.....	5
Ilustración 2. Diagrama general del algoritmo. Fuente: Propia.	6
Ilustración 3. Estrategia 1 (Búsqueda en Base de datos). Fuente: Autor.....	10
Ilustración 4. Estrategia 2 (Búsqueda por enlace roto en internet). Fuente: Autor.....	11
Ilustración 5. Estrategia 3 (Búsqueda por texto en Internet). Fuente: Autor.	12
Ilustración 6. Estrategia 4 (Búsqueda en archive.org y Google). Fuente: Autor.	13
Ilustración 7. Ejecución de las estrategias. Fuente: Autor.	14
Ilustración 8. Ing. M.Sc. Javier Enciso introduciendo la Aplicación BYE BYE 404. Fuente: Villadevs.	50
Ilustración 9. Pasante Maximiliano Cárdenas presentando forma de uso de BYE BYE 404. Fuente: Villadevs.	51
Ilustración 10. Pasante Emiraldo Lozano presentando algoritmo de recomendaciones implementado en BYE BYE 404. Fuente: Villadevs.....	51
Ilustración 11. Recomendación de IAU. Fuente: IAU.....	52



RESUMEN

El informe que se presenta a continuación tiene como fin contextualizar sobre los aspectos más importantes que se llevaron a cabo para la elaboración y puesta en marcha del ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB.

El lector podrá tener una visión panorámica del trabajo realizado en el desarrollo de este algoritmo, el cual tiene como objetivo recomendar la página equivalente a partir de un enlace roto. El algoritmo fue integrado en la aplicación web llamada BYE BYE 404. La cual es una aplicación que detecta enlaces rotos, los informa una vez encontrados y supervisa sitios web. Además, sugiere una posible solución para el enlace detectado.

JUSTIFICACIÓN

Según un estudio del 2013 citado por The New York Times [1], el 49 por ciento de los hipervínculos usados para sustentar decisiones de la Suprema Corte de los Estados Unidos, ya no funcionan; y lo que es aún peor, enlaces encontrados en el sitio web de la misma, que apuntaban a otras páginas dentro del mismo sitio, también fallan. En un experimento anterior, realizado en el 2003, Fetterly y otros, [2] descubrieron que aproximadamente 1 de cada 200 enlaces desapareció cada semana de Internet. Dos años después McCown y otros, descubrieron que la mitad de los documentos citados en la revista D-Lib ya no estaban disponibles después de 10 años de su publicación y otros estudios han demostrado que la decadencia de disponibilidad de los enlaces en la literatura académica es aún mayor [3].

Como se puede ver, mantener en funcionamiento todos los enlaces de un sitio web es una tarea complicada, que cada vez tiene mayor importancia, pues la información presente en Internet está siendo cada vez más usada como soporte en todo tipo de procesos académicos, comerciales e incluso judiciales, reemplazando las fuentes estáticas y permanentes usadas previamente: los libros.

Las diferentes instituciones han empezado a tomar medidas para intentar reparar los daños producidos por el no funcionamiento de los enlaces. En el caso de la corte, la solución por la que han optado consiste en tener registro de la fecha exacta en la que se visitó el sitio web, y guardar una copia física de lo consultado. Otras entidades estatales del país mantienen un archivo electrónico en formato PDF de lo que ellos llaman "webcites".

En el ámbito de las leyes, personas como Jonathan Zittrain, profesor de leyes y ciencias de la computación en Harvard, trabajan en soluciones más ambiciosas. En el caso de Zittrain es Perma.cc, una plataforma construida y mantenida por un consorcio de bibliotecas de leyes que



permitirá a escritores y editores capturar y guardar información normalmente transitoria en la web con un nuevo link, esta vez permanente [1].

OBJETIVOS

OBJETIVO GENERAL

- Diseñar e implementar un algoritmo que recomiende una página equivalente a partir de un enlace roto.

OBJETIVOS ESPECÍFICOS

- Diseñar e implementar distintas estrategias de búsqueda que permitan ofrecer una sugerencia.
- Integrar el algoritmo en la nueva versión de la aplicación BYE BYE 404.

MARCO DE REFERENCIA

La aplicación web BYE BYE 404, a petición de Mobile Corp S.A.S., se desplegó en una plataforma de Infraestructura como servicio (IaaS), fue desarrollada en lenguaje de programación Python con el framework de desarrollo web Django por ello el algoritmo se elaboró usando las siguientes tecnologías: UriLib, BeautifulSoup y Request. Las tecnologías anteriormente mencionadas están desarrolladas bajo el lenguaje de programación Python, además son compatibles con la versión de Django utilizada por el cliente.

En la Ilustración 1 se aprecia la arquitectura solución general del proyecto, donde se aprecia el lugar donde fue integrado el algoritmo desarrollado.

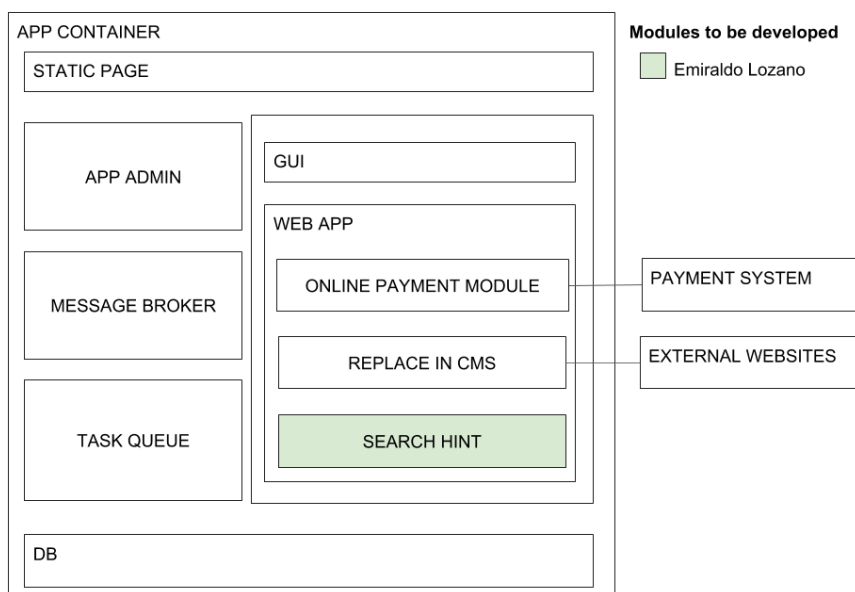


Ilustración 1, Arquitectura de Solución. Fuente: Propia.



En la Ilustración 2 se presenta de manera general el proceso que sigue el algoritmo para encontrar la mejor sugerencia.

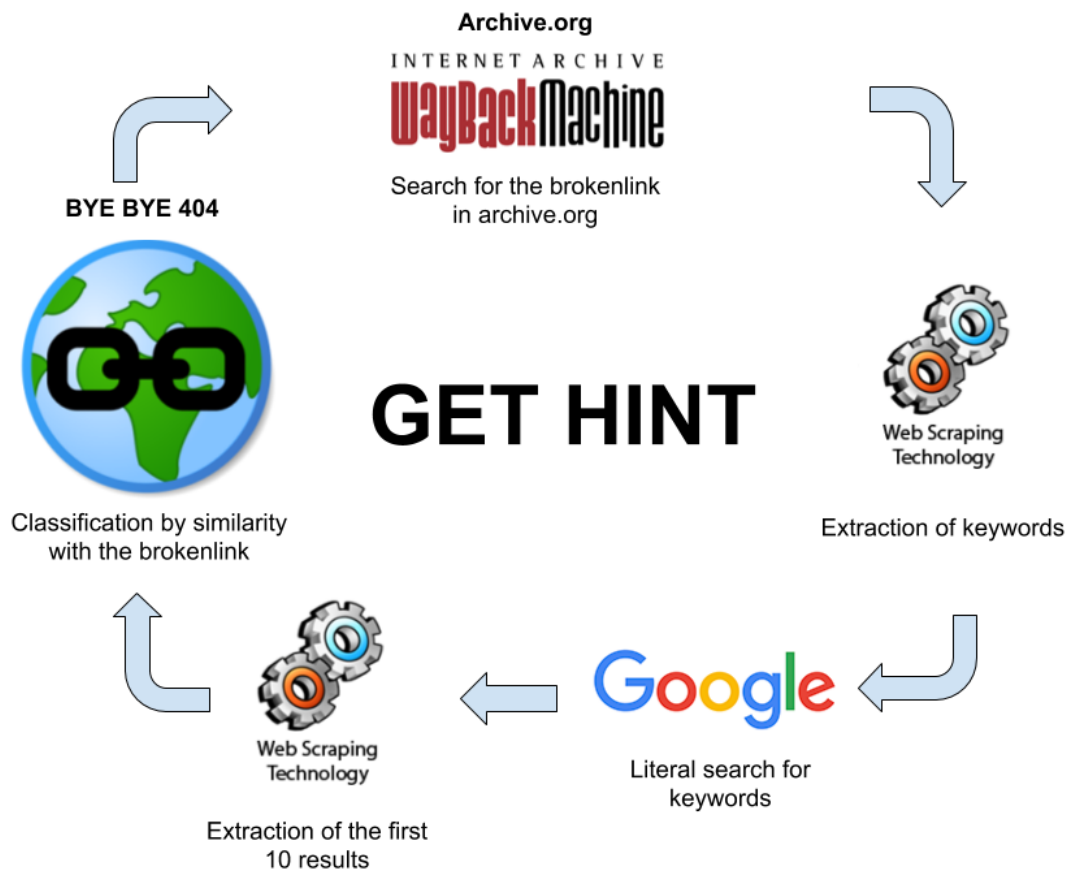


Ilustración 2. Diagrama general del algoritmo. Fuente: Propia.

REFERENCIA TEÓRICA

IaaS es una infraestructura informática inmediata que se aprovisiona y administra a través de Internet, permite reducir o escalar verticalmente los recursos con rapidez para ajustarlos a la demanda, se paga por uso evitando gastos innecesarios y la complejidad que supone la compra y administración de servidores físicos propios y todo el resto de la infraestructura de un centro de datos [4]. Por petición del cliente, se usará el potencial de la computación en la nube ofrecido por Amazon Web Services, específicamente los servicios a utilizar son Elastic Beanstalk, S3, RDS, EC2, SES, Route 53 [5].

Django es un framework de desarrollo web de código abierto escrito en Python. Permite construir aplicaciones web de forma más rápida, y con menos código, pues simplifica el trabajo de los programadores y hace que estos se concentren en codificar solo lo relativo al proyecto en específico [6]. Uno de los componentes más potentes de Django es su interfaz de administración automática, pues esta se encarga de leer los metadatos de los modelos para proporcionar una interfaz rápida centrada en el modelo en la que los usuarios autorizados pueden administrar el contenido de los mismos [7], [8], [9].



Python, a su vez, es un lenguaje de scripting multiplataforma que soporta distintos paradigmas de programación, entre los que se encuentran la Orientación a Objetos, con el que se puede crear cualquier tipo de programa e incluso páginas web. Es un lenguaje interpretado, por lo que tiene una menor velocidad de ejecución, aunque el tiempo de desarrollo sea reducido [10]. Django y Python fueron establecidos como requerimientos por el cliente.

Web Scraping es una técnica utilizada para extraer información de páginas web ya sea de forma manual o ejecutando scripts que realicen búsquedas de información que se considera atractiva o relevante. Una forma de realizar web scraping con lenguaje de programación Python es la implementación de la librería BeautifulSoup para el manejo sintáctico de algún archivo HTML extraído de una página web usando requests [11].

Archive.org o Internet Archive es una biblioteca digital cuyo objetivo principal es la preservación de archivos, sitios web y recursos multimedia. Fue creada en 1996 y actualmente posee un aproximado de 324 billones de capturas de sitios web [12].

BeautifulSoup es una librería para el lenguaje de programación Python, su función principal es extraer datos de archivos XML y HTML. Funciona con diferentes analizadores sintácticos tales como LXML o HTML5LIB para proporcionar diversas formas de navegación, búsqueda y modificación del árbol de análisis sintáctico [13].

Requests es una librería escrita en Python para realizar peticiones HTTP de forma sencilla con funciones como: Thread safety, connection pooling, la verificación SSL / TLS del lado del cliente, cargas de archivos con codificación multiparte, auxiliares para reintentar solicitudes y tratar con redireccionamientos HTTP, soporte para gzip y deflate encoding, Soporte proxy para HTTP y SOCKS [14].

TRABAJOS SIMILARES

- PageChaser: es un sistema que supervisa los enlaces que se encuentran dentro de una página web. Detecta cuando uno de estos enlaces se encuentra roto y busca la nueva ubicación de la página correspondiente partiendo de que el enlace ha dejado de funcionar porque fue movido de sección en la misma página web [15].
- Updating broken web links: es un trabajo de investigación centrado en la creación de un sistema para la recomendación de un enlace a una página web que sirva como reemplazo a un enlace roto, sea el caso de que encuentre la misma página web o que recomiende una muy similar a la que se busca inicialmente. Se emplean varios recursos de contexto tanto del enlace roto como de la página web que lo contiene para que la búsqueda sea más acertada tales como el texto que contiene la etiqueta del enlace roto, el texto que rodea al enlace roto dentro de la página web, la url de la página web que lo contiene, entre otros [16].



DESARROLLO

Recolección de Información

Búsqueda de plataformas con objetivos similares.

No se encontraron plataformas con objetivos similares, sin embargo, se han publicado trabajos similares al algoritmo desarrollado, tales como:

- PageChaser: A Tool for the Automatic Correction of Broken Web Links [15].
- Updating broken web links: An automatic recommendation system[16].

Análisis de los algoritmos de búsqueda y recuperación utilizados en las plataformas encontradas.

En los trabajos encontrados se detectan estrategias similares, las cuales consisten en:

- Extraer palabras claves del sitio web donde se encuentra el enlace roto
- Realizar consultas de estas en motores de búsqueda.
- Realizar análisis de similitud de los resultados respecto a la url original.

También, se evidencia el uso de una página caché para la extracción de palabras claves, además proponen la búsqueda del enlace roto en otras secciones del sitio web para verificar si existe aún dentro del mismo sitio web.

Definición y clasificación de estrategias de búsqueda.

A continuación, se pueden apreciar el planteamiento de las estrategias de búsqueda que se implementaron en el algoritmo propuesto y su clasificación por dificultad, basado en el número de tareas que debe realizar cada una para terminar su debido proceso.

Definición de estrategias de búsqueda.

1. Búsqueda en base de datos: Realizar una búsqueda en la base de datos (Base de datos de la plataforma BYE BYE 404) para corroborar si existen registros de búsquedas anteriores sobre el enlace roto, en caso de encontrar registros, verificar si se le ha asignado una solución y retornarla.
2. Búsqueda por texto: Realizar una búsqueda literal en el motor de búsqueda de Google [17] del texto que contiene la etiqueta HTML que referencia al enlace roto (Información suministrada por BYE BYE 404), evaluar entre los primeros 10 resultados cual tiene más semejanza a la url original y retornarlo.
3. Búsqueda por enlace roto: Realizar una búsqueda literal en el motor de búsqueda de Google del enlace roto, evaluar entre los primeros 10 resultados cual tiene más semejanza a la url original y retornarlo.



4. Búsqueda en archive.org y Google: Realizar una búsqueda en la página web <http://archive.org> del enlace roto, obtener el último registro (en caso de que exista) y guardar su código HTML, utilizar web scraping para la extracción de palabras u oraciones claves del código HTML, ejecutar una búsqueda literal por cada palabra u oración y clasificar los primeros 10 resultados por cantidad de repeticiones y semejanza a la url original, y finalmente retornar el enlace que más se repita o asemeje a la url original.
5. Búsqueda en archive.org: Realizar una búsqueda en la página web <http://archive.org> del enlace roto, obtener el último registro (en caso de que exista) y retornarlo.

Clasificación por complejidad de dichas estrategias.

1. Búsqueda en base de datos.
2. Búsqueda por enlace roto.
3. Búsqueda por texto.
4. Búsqueda en archive.org
5. Búsqueda en archive.org y Google.

Selección de tecnologías y estrategias a implementar.

Tecnologías:

Basados en la arquitectura de la aplicación BYE BYE 404 definida en la Ilustración 1 se define una lista de tecnologías a utilizar:

- Python 2.7.10
- BeautifulSoup 4.5.3
- Requests 2.9.1
- Google (Librería de terceros para realizar búsquedas de Google en lenguaje de programación Python) 1.9.3

Estrategias a implementar:

Seguidamente de analizar las estrategias se determinó que la estrategia “Búsqueda en archive.org” está implícita en la “Búsqueda en archive.org y Google” teniendo en cuenta que esta última al no determinar una sugerencia luego de hacer la búsqueda en Google, retornará el resultado encontrado en archive.org, dando como resultado las siguientes estrategias a implementar.

1. Búsqueda en base de datos.
2. Búsqueda por enlace roto.
3. Búsqueda por texto.
4. Búsqueda en archive.org y Google.



Diseño y desarrollo de estrategias

Diseño

En la Ilustración 3 se evidencia el diagrama de flujo de la estrategia “Búsqueda en base de datos”

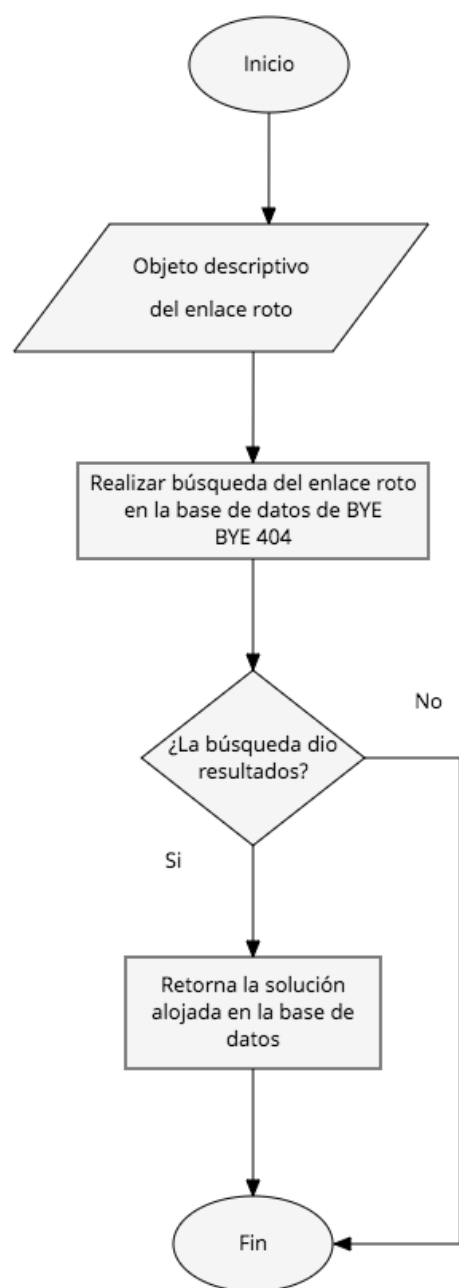


Ilustración 3. Estrategia 1 (Búsqueda en Base de datos). Fuente: Autor.



En la Ilustración 4 se evidencia el diagrama de flujo de la estrategia “Búsqueda por texto”

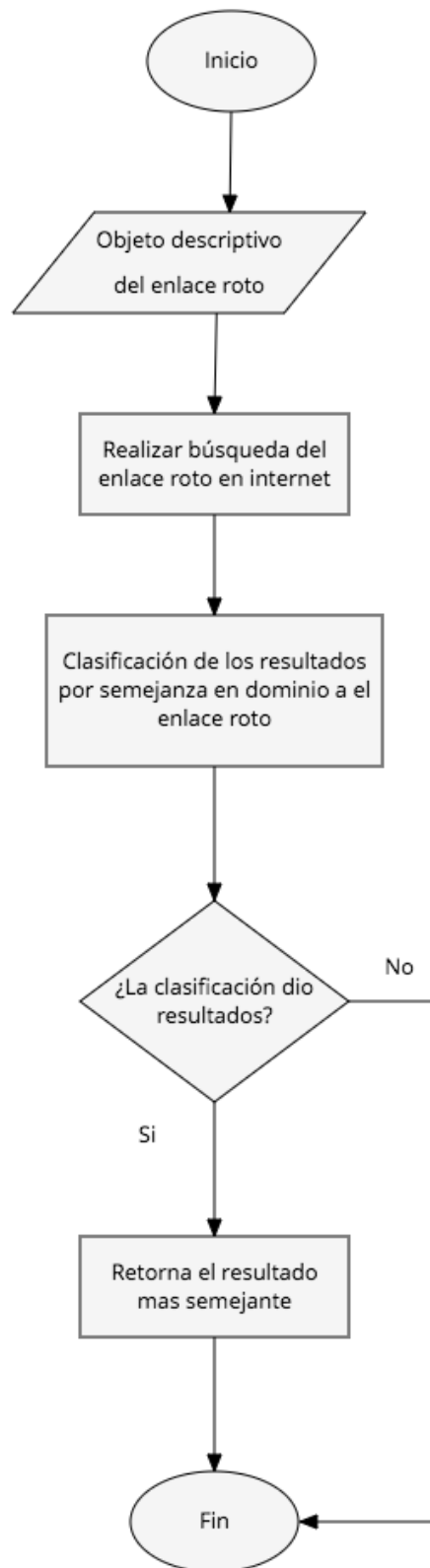


Ilustración 4. Estrategia 2 (Búsqueda por enlace roto en internet). Fuente: Autor.



En la Ilustración 5 se evidencia el diagrama de flujo de la estrategia “Búsqueda por enlace roto”

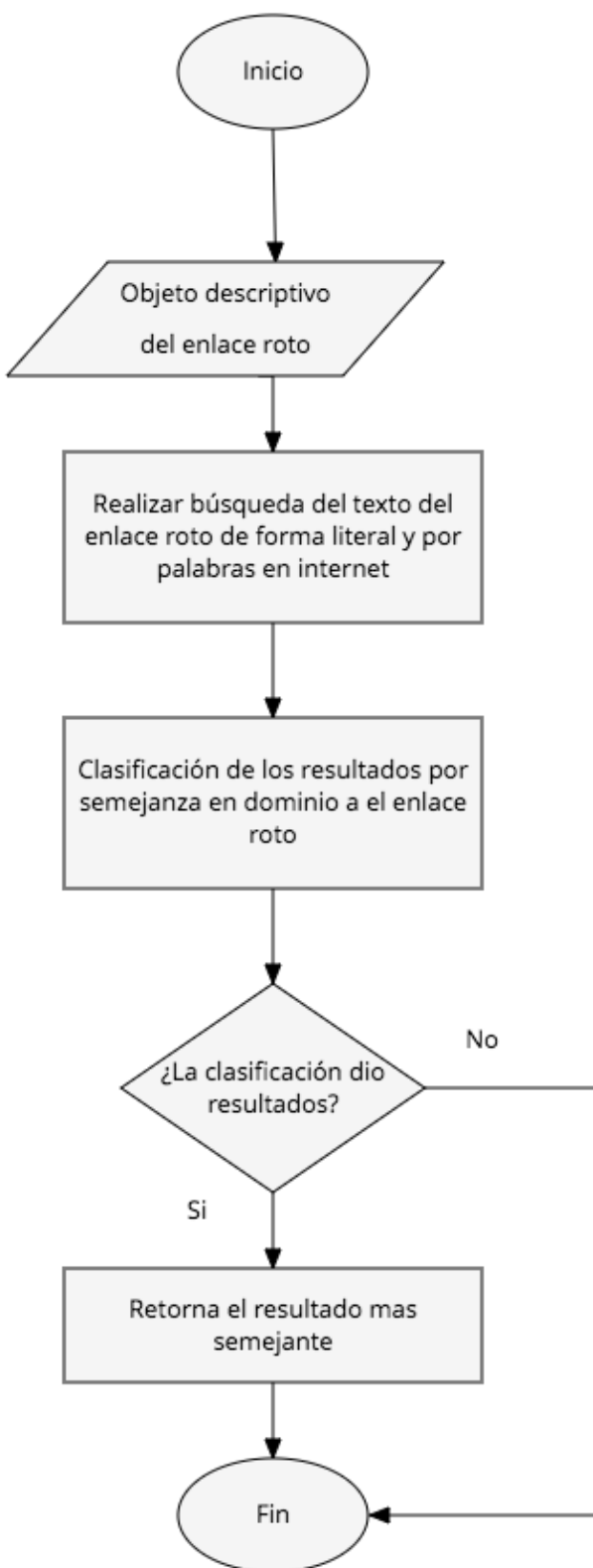


Ilustración 5. Estrategia 3 (Búsqueda por texto en Internet). Fuente: Autor.

En la Ilustración 6 se evidencia el diagrama de flujo de la estrategia “Búsqueda en archive.org y Google”

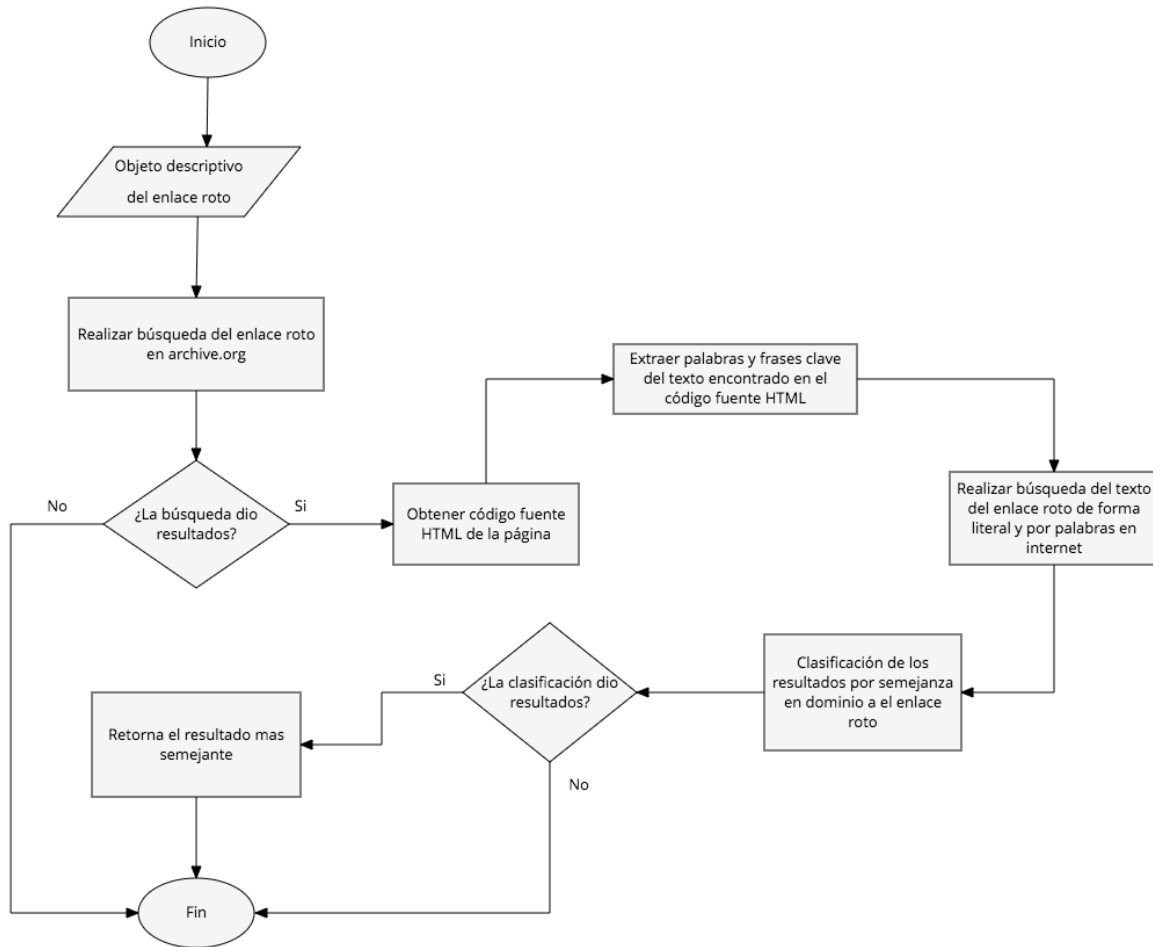


Ilustración 6. Estrategia 4 (Búsqueda en archive.org y Google). Fuente: Autor.

Orden de ejecución de estrategias

En la Ilustración 7 se evidencia el diagrama de flujo de ejecución de las estrategias.

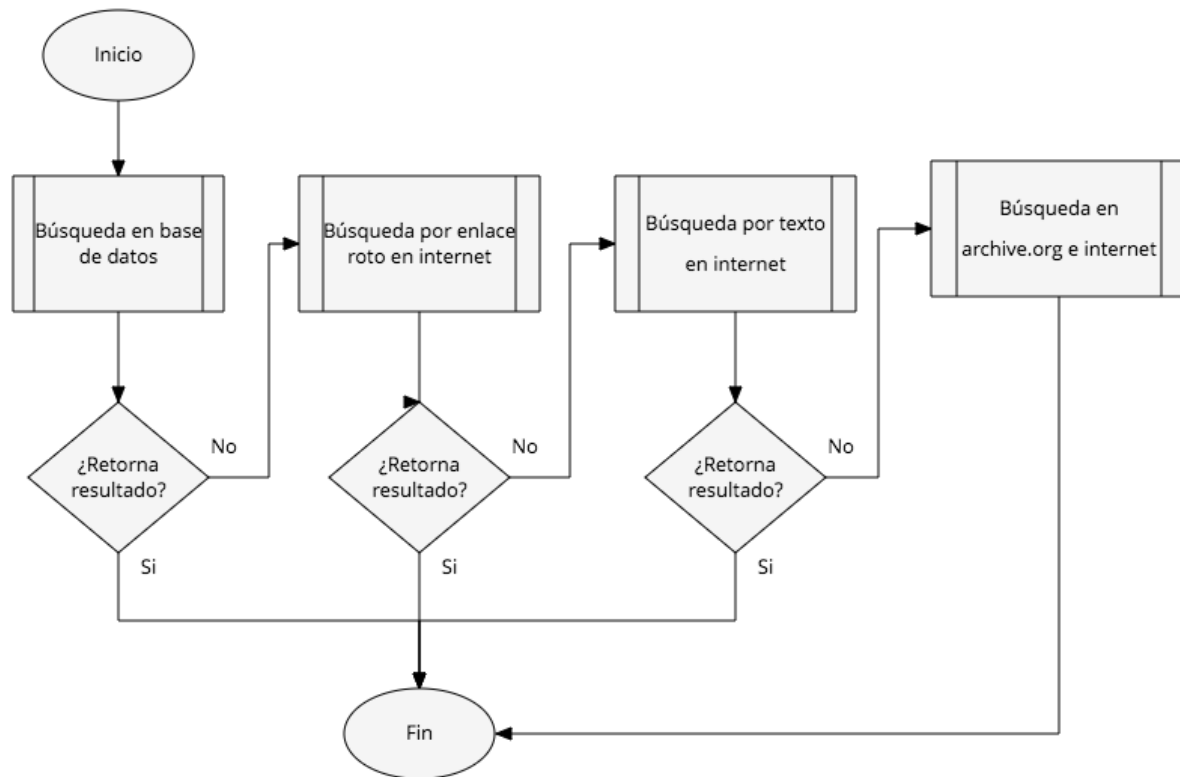


Ilustración 7. Ejecución de las estrategias. Fuente: Autor.

Desarrollo

Implementación de los algoritmos

En el APÉNDICE A. Algoritmo planteado. Se podrá evidenciar la implementación de los algoritmos y otras utilidades.

Dificultades encontradas en la implementación de estrategias

1. La base principal para el desarrollo de las estrategias propuestas es el uso de la técnica de extracción de información WebScraping. En esta, luego de hacer peticiones HTTP a una página web y obtener su código html, se realiza dicha extracción de este para posteriormente ser tratado y utilizado en la búsqueda de la mejor sugerencia reemplazando así el enlace roto; sin embargo, se encontraron dificultades con algunas páginas web en donde su contenido fue construido de tal forma que no se expresa como respuesta a la petición HTTP sobre el código HTML, sino a través de peticiones



asíncronas AJAX (Asynchronous JavaScript And XML) impidiendo así la extracción de información mediante el uso de WebScraping.

2. Durante el desarrollo de las estrategias planteadas se observó una tendencia en algunos tipos específicos de páginas web que dificultan la búsqueda de una sugerencia acertada. Se encontraron casos tales como: páginas web de eventos que tienden a desaparecer luego de que este se realiza, páginas web de noticias y educación que tienden a ser reubicadas, pasado un año, a otra porque se vuelven una traza histórica del sitio web.
3. Una de las mayores dificultades afrontadas en el desarrollo de las estrategias planteadas fueron los enlaces rotos denominados como falsos positivos, estos enlaces se presentan como rotos ante las estrategias y dan como resultado el mismo enlace. Una de las razones por las cuales puede ocurrir esto, es que la página que representa el enlace roto esté actualizando su registro de archivo robots.txt con el fin de evitar el uso de WebScraping para extraer su información; de esta manera, se observa como un enlace roto pero a la hora de realizar la búsqueda este mismo puede resultar como recomendación.

Desarrollo de ejecución de los algoritmos en Segundo plano.

En BYE BYE 404 la búsqueda de los enlaces rotos se realiza en una tarea programada en segundo plano. Al encontrar un enlace roto, BYE BYE 404 hará uso de los algoritmos propuestos al estar la búsqueda en una tarea en segundo plano se hace redundante realizar una ejecución en segundo plano de los algoritmos, pues estos al ser invocados en una tarea en segundo plano, no se ve reflejado en el tiempo de respuesta de las peticiones que se realizan normalmente a la plataforma.

Conexión con software empresarial

En el APÉNDICE B. Conexión con software empresarial. Se podrá evidenciar la implementación de los algoritmos planteados y funciones complementarias dentro del contexto de la aplicación BYE BYE 404, al estar la aplicación desarrollada con Django el algoritmo se trasladó al modelo que representa los enlaces rotos en la aplicación y su respectiva integración en las tareas programadas de búsqueda de enlaces rotos.

Comparación de rendimiento de las estrategias implementadas.

En estadística el tamaño de la muestra es el número de sujetos que componen la muestra extraída de una población, necesarios para que los datos obtenidos sean representativos de la población. La fórmula (1) se emplea para encontrar este dato [18].

$$n = \frac{N * \sigma^2 Z_{\alpha}^2}{e^2 (N-1) + \sigma^2 Z_{\alpha}^2} \quad (1)$$



Donde:

n = El tamaño de la muestra.

N = Tamaño de la población.

σ = Desviación estándar de la población, que generalmente cuando no se tiene su valor, suele utilizarse un valor estimado a ojo o a partir de una pequeña muestra o muestra piloto. Para ser conservador (prudente), mejor errar estimando por exceso que por defecto.

Z_α = Valor obtenido mediante niveles de confianza. Es un valor constante que, si no se tiene su valor, se lo toma en relación al 95% de confianza equivale a 1,64 (como más usual) o en relación al 99% de confianza equivale 2,33, valor que queda a criterio del encuestador.

e = Límite aceptable de error muestral que, generalmente cuando no se tiene su valor, suele utilizarse un valor que varía entre el 1% (0,01) y 9% (0,09), valor que queda a criterio del encuestador.

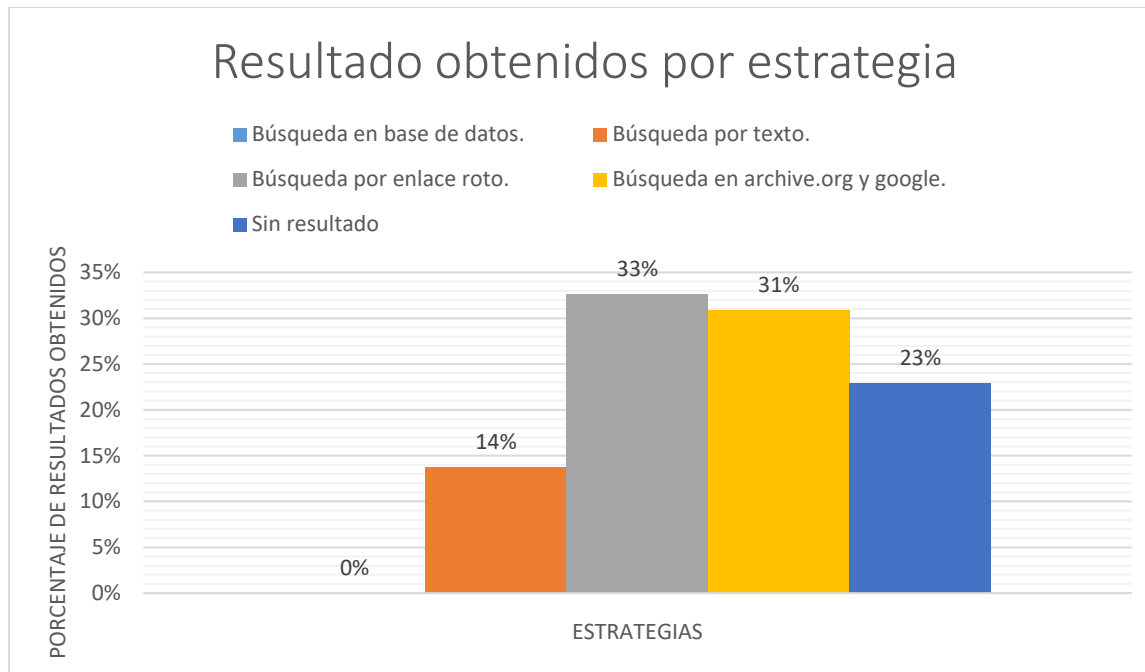
$$n = \frac{497 * 0.5^2 * 1.64^2}{0.05^2 * 496 + 0.5^2 * 1.64^2} \approx 174.74 \sim 175$$

Al desarrollar la formula (1), da como resultado 175, siendo esto la cantidad de registros que se tendrá en cuenta para medir la exactitud con la que el algoritmo planteado ofrece una solución.

Tabla 1. Comparación estrategias. Fuente: Autor.

Estrategia	Número de veces empleada la estrategia	Porcentaje registros encontrados por estrategia
Búsqueda por texto.	24	14%
Búsqueda en base de datos.	0	0%
Búsqueda por enlace roto.	57	33%
Búsqueda en archive.org y Google.	54	31%
Sin resultados.	40	23%

En la Tabla 1 se evidencian los resultados obtenidos al emplear cada una de las estrategias planteadas con 175 enlaces rotos de prueba aleatorios de los 497 basados en el tamaño de la muestra (formula 1). Se puede apreciar el total de resultados obtenidos en cada estrategia, ya que la ejecución de cada una de las estrategias es consecutiva y depende de que la estrategia anterior no obtenga resultado; evidenciamos que cada estrategia obtuvo un porcentaje de resultados con respecto al total de enlaces rotos de prueba usados, por ejemplo: de los 175 enlaces rotos de prueba usados, la estrategia “Búsqueda por texto” obtuvo un total de 24 resultados representando el 14% del total de enlaces rotos de prueba. Los registros se evidencian en APÉNDICE C.



Gráfica 1. Porcentaje de registros encontrados por estrategia. Fuente: Autor.

Definición de exactitud de la búsqueda que realiza el sistema.

Tabla 2. Efectividad búsqueda. Fuente: Autor.

Estrategia	Aciertos confirmados por estrategia	Porcentaje de aciertos por estrategia
Búsqueda por texto.	6	25%
Búsqueda en base de datos.	0	0%
Búsqueda por enlace roto.	32	56%
Búsqueda en archive.org y Google.	39	72%

En la Tabla 2 se presentan los resultados obtenidos al realizar un análisis humano de las recomendaciones obtenidas por cada una de las estrategias (valores previamente presentados en la Tabla 1). El procedimiento efectuado fue tomar cada uno de los valores encontrados por las estrategias y verificar manualmente haciendo una búsqueda en internet para corroborar que la recomendación dada por la estrategia es la más acertada, de este modo se obtuvo un porcentaje de efectividad de búsqueda por cada estrategia. Por ejemplo, para la estrategia “Búsqueda por texto” se obtuvo un valor de 25% de efectividad ya que sólo 6 recomendaciones de las 24 encontradas previamente, fueron las más acertadas.



RESULTADOS OBTENIDOS

- En la Tabla 3 del APÉNDICE C. Evidencia resultados. se presentan sólo los primeros 150 registros del muestreo. En la Tabla 2 se presentó un resumen de un total de 497 registros.
- La implementación de la estrategia “Búsqueda en base de datos” se ve reflejada en el APÉNDICE A. Algoritmo planteado. Con el nombre de `get_hint_db` y su integración con la aplicación se ve reflejada en el APÉNDICE B. Conexión con software empresarial. Con el mismo nombre.
- La implementación de la estrategia “Búsqueda por texto” se ve reflejada en el APÉNDICE A. Algoritmo planteado. Con el nombre de `get_hint_from_text` y su integración con la aplicación se ve reflejada en el APÉNDICE B. Conexión con software empresarial. Con el mismo nombre.
- La implementación de la estrategia “Búsqueda por enlace roto” se ve reflejada en el APÉNDICE A. Algoritmo planteado. Con el nombre de `get_hint_broken_url` y su integración con la aplicación se ve reflejada en el APÉNDICE B. Conexión con software empresarial. Con el mismo nombre.
- La implementación de la estrategia “Búsqueda en archive.org y Google” se ve reflejada en el APÉNDICE A. Algoritmo planteado. Con el nombre de `get_hint_search_in_archive` y su integración con la aplicación se ve reflejada en el APÉNDICE B. Conexión con software empresarial. Con el mismo nombre.

CONCLUSIONES

- Web Scraping es una técnica poderosa a la hora de extraer información de páginas web o archivos de hipertexto. Sin embargo, Web Scraping presenta limitaciones en las páginas web desarrolladas con frameworks tales como: AngularJS, Polymer, entre otros, ya que estos cargan su contenido a través de peticiones Ajax.
- Omitir un texto descriptivo en las etiquetas `<a>` es una mala práctica de los editores de contenido web que impide que la estrategia “Búsqueda por texto” produzca resultados satisfactorios.
- Dentro de las sugerencias propuestas por las estrategias se evidencian casos especiales en los que la sugerencia es el mismo “enlace roto”, estos enlaces fueron clasificados como falsos positivos y no se tuvieron en cuenta como solución acertada. Esto debido a que al acceder a enlaces externos del sitio web algunos presentan seguridad anti-robots, generando así un falso enlace roto que posteriormente realizando una búsqueda en internet del mismo el resultado más favorable va ser el mismo “enlace roto”, dando así una solución “acertada”.
- Dependiendo el contenido de la página web se observan tendencias tales como:
 - Las páginas de eventos suelen desaparecer después de finalizado el evento.



- Las páginas web institucionales, gubernamentales, educativas, noticiosas o informativas suelen ser reubicadas al pasar de un año a otro.
- Las páginas web personales cambian con mayor frecuencia de dominio, pero su contenido no varía mucho.

RECOMENDACIONES

- Implementar la distancia de Levenshtein en la comparación por semejanza para evitar resultados nulos.
- Cuando existen resultados semejantes se debería realizar una búsqueda semántica partiendo del contenido de los posibles enlaces solución, a comparación con el resultado de la búsqueda en archive del enlace roto.
- Incluir diferentes buscadores tales como: Bing, DuckDuckGo, Yahoo!, entre otros, para obtener diversos resultados y evitar bloqueos por búsqueda recurrente.
- Crear una Base de Datos y almacenar palabras y textos claves de los enlaces activos de las páginas de los usuarios de BYE BYE 404, de esta forma cuando un enlace se registre como roto se tiene información necesaria para hacer una búsqueda y encontrar su posible solución o reemplazo.

REFERENCIAS

- [1] A. Liptak, "In Supreme Court Opinions, Web Links to Nowhere," 2013.
- [2] D. Fetterly, M. Manasse, M. Najork, and J. Wiener, "A large-scale study of the evolution of web pages," in *Proceedings of the twelfth international conference on World Wide Web - WWW '03*, 2003, p. 669.
- [3] F. McCown, C. Sheffan, M. L. Nelson, and J. Bollen, "The Availability and Persistence of Web References in D-Lib Magazine," in *5th International Web Archiving Workshop (TAW05)*, 2005.
- [4] Mateo Santos, "SaaS, IaaS y PaaS: ¿qué son, cómo usarlos y para qué? • ENTER.CO," *enter.co*, 2015. [Online]. Available: <http://www.enter.co/guias/tecnoguias-para-empresas/saas-iaas-y-paas-que-son-como-usarlos-y-para-que/>. [Accessed: 22-Nov-2017].
- [5] Amazon Web Services Inc, "¿Qué es AWS? - Amazon Web Services," *amazon.com*, 2017. [Online]. Available: <https://aws.amazon.com/es/what-is-aws/>. [Accessed: 04-Jun-2017].
- [6] djangoproject.com, "The Web framework for perfectionists with deadlines | Django." [Online]. Available: <https://www.djangoproject.com/>. [Accessed: 03-Jun-2017].
- [7] DjangoGirls, "¿Qué es Django?," *djangogirls.org*, 2015. [Online]. Available: <https://tutorial.djangogirls.org/es/django/>. [Accessed: 14-Aug-2017].
- [8] J. R. Molina Ríos, N. M. Loja Mora, M. P. Zea Ordóñez, and E. L. Loaiza Sojos, "Evaluación de los Frameworks en el Desarrollo de Aplicaciones Web con Python," *Rev. Latinoam. Ing. Softw.*, vol. 4, no. 4, p. 201, Sep. 2016.
- [9] Django Software Foundation, "The Django admin site | Django documentation | Django," *Django Software Foundation*, 2017. [Online]. Available:



- <https://docs.djangoproject.com/en/1.9/ref/contrib/admin/>. [Accessed: 11-Jun-2017].
- [10] A. Sweigart, *Automate the boring stuff with Python: practical programming for total beginners*. .
 - [11] A. T. Hernández, E. Gómez Vázquez, A. B. Rincón, J. Montero García, A. C. Maldonado, and R. Ibarra Orozco, “Metodologías para análisis político utilizando Web Scraping,” *Res. Comput. Sci.*, vol. 95, pp. 113–121, 2015.
 - [12] archive.org, “Internet Archive: About IA,” *archive.org*, 2017. [Online]. Available: <https://archive.org/about/>. [Accessed: 22-Nov-2017].
 - [13] BBVAOpen4U, “Herramientas de extracción de datos: para principiantes y profesionales | BBVAOpen4u.com. Actualidad y eventos sobre APIS, Fintech y Open Banking,” *bbvaopen4u.com*, 2016. [Online]. Available: <https://bbvaopen4u.com/es/actualidad/herramientas-de-extraccion-de-datos-para-principiantes-y-profesionales>. [Accessed: 22-Nov-2017].
 - [14] urllib3, “urllib3 — urllib3 dev documentation,” *readthedocs.io*, 2017. [Online]. Available: <https://urllib3.readthedocs.io/en/latest/>. [Accessed: 22-Nov-2017].
 - [15] A. Morishima, A. Nakamizo, T. Iida, S. Sugimoto, and H. Kitagawa, “PageChaser: A Tool for the Automatic Correction of Broken Web Links,” in *2008 IEEE 24th International Conference on Data Engineering*, 2008, pp. 1486–1488.
 - [16] J. Martinez-Romo and L. Araujo, “Updating broken web links: An automatic recommendation system,” *Inf. Process. Manag.*, vol. 48, no. 2, pp. 183–203, Mar. 2012.
 - [17] google.com, “Search Tips & Tricks – Inside Search – Google,” *google.com*, 2016. [Online]. Available: https://www.google.com/intl/en_u/insidesearch/tipstricks/all.html. [Accessed: 18-Feb-2018].
 - [18] P. I. M. Torres, I. K. Paz, I. Federico, and G. Salazar, “TAMAÑO DE UNA MUESTRA PARA UNA INVESTIGACIÓN DE MERCADO.”



APÉNDICE A. Algoritmo planteado.

```
import csv
import urllib2
import urlparse
import requests
from bs4 import BeautifulSoup
from Google import search

def scraping_label_search_Google(labels, html, limit):
    """
    Scraping on html code, search a param label. Also search in Google the
    results.
    :param label: [] names the labels to search in html, example 'h1'
    :param html: html code to search the label, type BeautifulSoup in parse
    'lxml'
    :param limit: is the limit of results in 'label' search
    :return: the url's found in Google [priority, others]
    """
    others = []
    priority = []
    for x in labels:
        texts = html.find_all(x, )
        for text in texts[:limit]:
            keyword = text.get_text(strip=True)
            if len(keyword.split('.')) > 0:
                keyword = keyword.split('.')
                for key in keyword:
                    count = 0
                    try:
                        rt = search('"' + key.encode('ascii', 'ignore') + '"',
stop=10)
                    except:
                        print "Google blocks the connection, results are not
output."

                        rt = []
                        for url in rt:
                            others.append(url)
                            count += 1
                        if count == 1:
                            priority.append(others[-1])
                            others.pop()
    return priority, others

def get_domain_url(url):
    """
    return de domain from an url
    :param url:url to extract domain
    :return:the domain example : url = www.Google.com, return Google
    """
    # if is not full url, return None, example for relative urls.
    # Structure accepted:
    # http://www.example.com
    # http://www.example.com/home/
    # http://web.example.hello.com
```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```
# http://web.example.hello.com/home/
try:
    url = url
    url1 = url.split('//')[1]
    url2 = url1.split('/')[0]
    url3 = url2.split('.')
    url3 = url3[1:len(url3) - 1]
    url4 = ""
    if len(url3) > 1:
        for i in url3:
            url4 += i + "."
        url4 = url4[:len(url4) - 1]
    else:
        for i in url3:
            url4 += i
    return url4
except:
    return None

def get_possible_url(url, results):
    """
    Returns the possible url by analyzing for the domain
    :param url: brake url
    :param results: resutls search in Google
    :return: possible url
    """
    results = results
    possible = []
    url1 = get_domain_url(url)
    for i in results:
        if get_domain_url(i) == url1:
            possible.append(i)
    return possible

def get_hint_db(link):
    hint = ''
    hints = open("hints.csv")
    input_hints = csv.DictReader(hints)
    for row in input_hints:
        row['broken_url']
        if row['broken_url'] == link:
            hint = row['hint_url']
            break
    hints.close()
    return hint

def get_hint_from_text(url, text):
    results = []
    keywords = text.split(' ')
    for text in keywords:
        keyword = text

        try:
            for url in search(keyword, stop=10):
```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```
        results.append(url)
    except:
        pass

    try:
        for url in search(text, stop=10):
            results.append(url)
    except:
        pass
    possible = get_possible_url(url=url, results=results)
    return possible

def get_hint_broken_url(url):
    results = []
    for response in search('"' + url + '"', stop=10):
        results.append(response)
    possible = get_possible_url(url=url, results=results)
    return possible

def get_hint_search_in_archive(url):
    hint = ''
    full_url = url
    url_archive = 'http://archive.org/wayback/available?url='
    rta_archive = requests.get(url_archive + url + '&timestamp=*', )
    last_working_page_url = ''
    if rta_archive.json()['archived_snapshots']:
        # url last working page
        last_working_page_url =
    rta_archive.json()['archived_snapshots']['closest']['url']
    page = ''
    try:
        last_working_page_url =
    'http://' + last_working_page_url.split('http://')[1][:-
1] + 'if_/http://' + last_working_page_url.split('http://')[2]
    page = urllib2.urlopen(last_working_page_url).read()
    except:
        print "Error, not open page."
    soup = BeautifulSoup(page, 'lxml')
    if soup.find_all('frame', ):
        # if have frame
        texts = soup.find_all('frame', )
        src = []
        for x in texts:
            src.append(x.get('src'))
        for x in src:
            # Scroll through the frames found
            page = None
            # Classify if it is a file or url
            if x.find('/') != -1:
                # Obtain the html of the url
                page = urllib2.urlopen('http://web.archive.org' +
str(x)).read()
            else:
                # Obtain the url of the file
                if last_working_page_url[-1] != '/':
```



```

        change = last_working_page_url
        change = change.replace(change.split('/')[ -1], x)
        page = urllib2.urlopen(change).read()
    else:
        page = urllib2.urlopen(last_working_page_url +
str(x)).read()
    soup = BeautifulSoup(page, 'lxml')

    priority, others = scraping_label_search_Google(labels=['p',
], html=soup, limit=6)
    print priority
    print others
    # Rta is the dict that keeps repeated of priority in others
    rt = {}
    for x in priority:
        for y in others:
            if x == y:
                try:
                    rt.update({x: rt[x] + 1})
                except:
                    rt.update({x: 1})
    if len(rt) == 1:
        url_rt, repeat = rt.items()[0]
        return url_rt
    else:
        # When there is more than one url with a maximum of
repetition in rta
        for x in others:
            rt.update({x: others.count(x)})
        url = rt.keys()
        repeat = rt.values()
        if url and repeat:
            max_repeat = max(repeat)
            rt_max_repeat = []
            for x in range(0, len(url) - 1):
                if repeat[x] == max_repeat:
                    rt_max_repeat.append(url[x])
            possible_url = get_possible_url(url=url,
results=rt_max_repeat)
            return possible_url
        else:
            priority, others = scraping_label_search_Google(labels=['h1',
'p', 'h2'], html=soup, limit=6)
            # Rta is the dict that keeps repeated of priority in others
            rt = {}
            for x in priority:
                for y in others:
                    if x == y:
                        try:
                            rt.update({x: rt[x] + 1})
                        except:
                            rt.update({x: 1})
            if len(rt) == 1:
                url_rt, repeat = rt.items()[0]
                return url_rt
            else:
                for x in others:

```



```
        rt.update({x: others.count(x)})
url_rt = rt.keys()
repeat = rt.values()
if url_rt and repeat:
    max_repeat = max(repeat)
    rt_max_repeat = []
    for x in range(0, len(url) - 1):
        if repeat[x] == max_repeat:
            rt_max_repeat.append(url[x])
    possible_url = get_possible_url(url=url,
results=rt_max_repeat)
    if possible_url:
        return possible_url
    else:
        # Did not find a hint
        return last_working_page_url
else:
    # Not found page in archive.org
    last_working_page_url = False
return hint
```

APÉNDICE B. Conexión con software empresarial.

```
def get_domain_url(self, url):
    """
    return de domain from an url
    :param url:url to extract domain
    :return:the domain example : url = www.Google.com, return Google
    """
    # if is not full url, return None, example for relative urls.
    # Structure accepted:
    # http://www.example.com
    # http://www.example.com/home/
    # http://web.example.hello.com
    # http://web.example.hello.com/home/
    try:
        url = url
        url1 = url.split('//')[1]
        url2 = url1.split('/')[0]
        url3 = url2.split('.')
        url3 = url3[1:len(url3) - 1]
        url4 = ""

        if len(url3) > 1:
            for i in url3:
                url4 += i + "."
            url4 = url4[:len(url4) - 1]
        else:
            for i in url3:
                url4 += i

        return url4
    except:
        return None
```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```
def get_possible_url(self, url, results):
    """
    Returns the possible url by analyzing for the domain
    :param url: brake url
    :param results: results search in Google
    :return: possible url
    """
    results = results
    possible = []

    url1 = self.get_domain_url(url)

    for i in results:

        if self.get_domain_url(i) == url1:
            possible.append(i)

    return possible

def scraping_label_search_Google(self, labels, html, limit):
    """
    Scraping on html code, search a param label. Also search in Google the
    results.
    :param label: [] names the labels to search in html, example 'h1'
    :param html: html code to search the label, type BeautifulSoup in parse
    'lxml'
    :param limit: is the limit of results in 'label' search
    :return: the url's found in Google [priority, others]
    """
    others = []
    priority = []
    for x in labels:
        texts = html.find_all(x, )
        for text in texts[:limit]:
            keyword = text.get_text(strip=True)
            if len(keyword.split('.')) > 0:
                keyword = keyword.split('.')
            for key in keyword:
                print "key" + key
                count = 0
                try:
                    rt = search('"' + key.encode('ascii', 'ignore') + '"',
stop=10)

                except:
                    print "Google blocks the connection, results are not
output."

                    rt = []
                    for url in rt:
                        others.append(url)
                        count += 1
                    if count == 1:
                        priority.append(others[-1])
                        others.pop()

    return priority, others

def get_hint(self):
```



```

"""
1. get full url of the broken link
2. search on archive.org
3. get the latest working page
4. extract the page title/first sentence
5. search (on search engine) for the new page using keywords from 4.
6. return the new url (if url internal, return realtive url)
"""
hint = ''
print "Apply the 1 strategy"
hint = self.get_hint_db()
if hint:
    return hint

print "Apply the 2 strategy"
hint = self.get_hint_broken_url()
if hint:
    return hint

print "Apply the 3 strategy"
hint = self.get_hint_from_text()
if hint:
    return hint

print "Apply the 4 strategy"
hint = self.get_hint_search_in_archive()
if hint.startswith(u'Not') or str(hint) == '':
    self.updated = datetime.now()
    self.save()
return hint

def get_hint_db(self):
    """
    Search the hint in the database and create a new register in model
    "hint"
    """
    hint = Hint.objects.filter(broken_url=self.url).first()
    if hint:
        print "option 1"
        # Hint(website=self.page.website, broken_url=hint.broken_url,
        fixed_url=hint.fixed_url).save()
        self.hint_proxy = hint.fixed_url
        self.save()
        hint.save()
        hint.rank = 3
        hint.save()
        self.upate_all_hint()
        return hint.fixed_url
    return None

def get_hint_from_text(self):
    """
    Search the hint based in the text of the attribute "text" in the object
    "link"
    """
    if self.text:
        results = []

```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```
keywords = self.text.split(' ')
for text in keywords:
    keyword = text

    try:
        for url in search(keyword, stop=10):
            results.append(url)
    except:
        pass

    try:
        for url in search(self.text, stop=10):
            results.append(url)
    except:
        pass

possible = self.get_possible_url(url=self.url, results=results)

print "the possible url"
print possible
if len(possible) > 0:
    if len(possible) >= 3:
        hint = Hint(website=self.page.website,
broken_url=self.url, fixed_url=possible[0], rank=3,
                    hint_origin='txt')
        hint.save()
        hint = Hint(website=self.page.website,
broken_url=self.url, fixed_url=possible[1], rank=2,
                    hint_origin='txt')
        hint.save()
        hint = Hint(website=self.page.website,
broken_url=self.url, fixed_url=possible[2], rank=1,
                    hint_origin='txt')
        hint.save()
    else:
        hint = Hint(website=self.page.website,
broken_url=self.url, fixed_url=possible[0], rank=3,
                    hint_origin='txt')
        hint.save()
        self.hint_proxy = possible[0]
        self.save()
        self.upate_all_hint()
        return possible[0]

def get_hint_broken_url(self):
    """
    Search the broken url in Google and select the hint
    """
    results = []
    full_url = self.get_full_url()
    for url in search('"' + full_url + '"', stop=10):
        results.append(url)

    possible = self.get_possible_url(url=full_url, results=results)

    if len(possible) > 0:
        if len(possible) >= 3:
```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```
        hint = Hint(website=self.page.website, broken_url=self.url,
fixed_url=possible[0], rank=3,
                    hint_origin='url')
        hint.save()

        hint = Hint(website=self.page.website, broken_url=self.url,
fixed_url=possible[1], rank=2,
                    hint_origin='url')
        hint.save()
        hint = Hint(website=self.page.website, broken_url=self.url,
fixed_url=possible[2], rank=1,
                    hint_origin='url')
        hint.save()
    else:
        hint = Hint(website=self.page.website, broken_url=self.url,
fixed_url=possible[0], rank=3,
                    hint_origin='url')
        hint.save()
        self.hint_proxy = possible[0]
        self.save()
        self.upate_all_hint()
        return possible[0]

def get_hint_search_in_archive(self):
    """
    Search the broken url in archive.org, read the results and search en
    Google coincidences
    """
    hint = ''

    full_url = self.get_full_url()

    url_archive = 'http://archive.org/wayback/available?url='
    rta_archive = requests.get(url_archive + full_url + '&timestamp=*', )
    last_working_page_url = ''
    if rta_archive.json()['archived_snapshots']:
        # url last working page
        last_working_page_url =
rta_archive.json()['archived_snapshots']['closest']['url']
        print last_working_page_url
        page = ''
        try:
            page = urllib2.urlopen(last_working_page_url).read()
        except:
            print "Error, not open page."
            soup = BeautifulSoup(page, 'lxml')
            if soup.find_all('frame', ):
                # if have frame
                texts = soup.find_all('frame', )
                src = []
                for x in texts:
                    src.append(x.get('src'))

                for x in src:
                    # Scroll through the frames found
                    page = None
                    # Classify if it is a file or url
```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```
if x.find('/') != -1:
    # Obtain the html of the url
    page = urllib2.urlopen('http://web.archive.org' +
str(x)).read()
else:
    # Obtain the url of the file
    if last_working_page_url[-1] != '/':
        change = last_working_page_url
        change = change.replace(change.split('/')[-1], x)
        page = urllib2.urlopen(change).read()
    else:
        page = urllib2.urlopen(last_working_page_url +
str(x)).read()

soup = BeautifulSoup(page, 'lxml')

priority, others =
self.scraping_label_search_Google(labels=['p', ], html=soup, limit=6)
# Rta is the dict that keeps repeated of priority in others
rt = {}
for x in priority:
    for y in others:
        if x == y:
            try:
                rt.update({x: rt[x] + 1})
            except:
                rt.update({x: 1})
if len(rt) == 1:
    url, repeat = rt.items()[0]
    Hint(website=self.page.website, broken_url=self.url,
        fixed_url=url, rank=3).save()
    self.hint_proxy = url
    self.save()
    self.upate_all_hint()
    return url
else:
    # When there is more than one url with a maximum of
repetition in rta
    for x in others:
        rt.update({x: others.count(x)})
    url = rt.keys()
    repeat = rt.values()
    if url and repeat:
        max_repeat = max(repeat)
        rt_max_repeat = []
        for x in range(0, len(url) - 1):
            if repeat[x] == max_repeat:
                rt_max_repeat.append(url[x])

    possible_url = self.get_possible_url(url=full_url,
results=rt_max_repeat)
    print possible_url
    if len(possible_url) >= 3:
        Hint(website=self.page.website,
broken_url=self.url,
            fixed_url=possible_url[0], rank=3,
hint_origin='archive_search').save()
```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```

                                Hint(website=self.page.website,
broken_url=self.url,
                                fixed_url=possible_url[1], rank=2,
hint_origin='archive_search').save()
                                Hint(website=self.page.website,
broken_url=self.url,
                                fixed_url=possible_url[2], rank=1,
hint_origin='archive_search').save()
                                self.hint_proxy = possible_url[0]
                                self.save()
                                self.upate_all_hint()
                                return possible_url[0]
                                elif len(possible_url) > 2:
                                Hint(website=self.page.website,
broken_url=self.url,
                                fixed_url=possible_url[0], rank=3,
hint_origin='archive_search').save()
                                Hint(website=self.page.website,
broken_url=self.url,
                                fixed_url=possible_url[1], rank=2,
hint_origin='archive_search').save()
                                self.hint_proxy = possible_url[0]
                                self.save()
                                self.upate_all_hint()
                                return possible_url[0]
                                elif len(possible_url) == 1:
                                Hint(website=self.page.website,
broken_url=self.url,
                                fixed_url=possible_url[0], rank=3,
hint_origin='archive_search').save()
                                self.hint_proxy = possible_url[0]
                                self.save()
                                self.upate_all_hint()
                                return possible_url[0]
                                else:
                                # Did not find a hint
                                Hint(website=self.page.website,
broken_url=self.url,
                                fixed_url=last_working_page_url, rank=3,
hint_origin='archive').save()
                                self.hint_proxy = last_working_page_url
                                self.save()
                                self.upate_all_hint()
                                return last_working_page_url

                                else:
                                priority, others =
self.scraping_label_search_Google(labels=['h1', 'p', 'h2'], html=soup,
limit=6)
                                # Rta is the dict that keeps repeated of priority in others
                                rt = {}
                                for x in priority:
                                for y in others:
                                if x == y:
                                try:
                                rt.update({x: rt[x] + 1})
                                except:
```

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



```
rt.update({x: 1})

if len(rt) == 1:
    url, repeat = rt.items()[0]
    Hint(website=self.page.website, broken_url=self.url,
          fixed_url=url, rank=3,
hint_origin='archive_search').save()
    self.hint_proxy = url
    self.save()
    self.upate_all_hint()
    return url
else:
    for x in others:
        rt.update({x: others.count(x)})
    url = rt.keys()
    repeat = rt.values()
    print repeat
    print url
    if url and repeat:
        max_repeat = max(repeat)
        rt_max_repeat = []
        for x in range(0, len(url) - 1):
            if repeat[x] == max_repeat:
                rt_max_repeat.append(url[x])

        print rt_max_repeat

    # Description of line code

    possible_url = self.get_possible_url(url=full_url,
results=rt_max_repeat)
    print possible_url
    if len(possible_url) >= 3:
        Hint(website=self.page.website, broken_url=self.url,
              fixed_url=possible_url[0], rank=3,
hint_origin='archive_search').save()
        Hint(website=self.page.website, broken_url=self.url,
              fixed_url=possible_url[1], rank=2,
hint_origin='archive_search').save()
        Hint(website=self.page.website, broken_url=self.url,
              fixed_url=possible_url[2], rank=1,
hint_origin='archive_search').save()
        self.hint_proxy = possible_url[0]
        self.save()
        self.upate_all_hint()
        return possible_url[0]
    elif len(possible_url) > 2:
        Hint(website=self.page.website, broken_url=self.url,
              fixed_url=possible_url[0], rank=3,
hint_origin='archive_search').save()
        Hint(website=self.page.website, broken_url=self.url,
              fixed_url=possible_url[1], rank=2,
hint_origin='archive_search').save()
        self.hint_proxy = possible_url[0]
        self.save()
        self.upate_all_hint()
```



```
        return possible_url[0]
    elif len(possible_url) == 1:
        Hint(website=self.page.website, broken_url=self.url,
              fixed_url=possible_url[0], rank=3,
hint_origin='archive_search').save()
        self.hint_proxy = possible_url[0]
        self.save()
        self.upate_all_hint()
        return possible_url[0]
    else:
        # Did not find a hint
        Hint(website=self.page.website, broken_url=self.url,
              fixed_url=last_working_page_url, rank=3,
hint_origin='archive').save()
        self.hint_proxy = last_working_page_url
        self.save()
        self.upate_all_hint()
        return last_working_page_url
else:
    # Not found page in archive.org
    last_working_page_url = False

    return "Not found possible url."

return hint
```

**APÉNDICE C. Evidencia resultados.***Tabla 3. Evidencia resultados obtenidos con algoritmo. Fuente: Autor.*

broken_link	text	hint	page
http://almaobservatory.org/en/press-room/press-releases/666-alma-discovers-a-formation-site-of-a-giant-planetary-system	here		https://www.eso.org/public/news/eso1549/
http://apps.microsoft.com/windows/en-us/app/european-southern-observatory/685c56a9-9b4d-4301-9de5-e67459638ba2	Windows 8.1 tablets and PCs	https://www.microsoft.com/en-us/store/p/european-southern-observatory/9wzdnrcdf64k	https://www.eso.org/public/announcements/ann14001/
http://astro.uchicago.edu/~kbs/wasp43b.html	video		https://www.spacetelescope.org/news/heic1422/
http://b612foundation.org/so-how-big-are-these-neos-anyway/	b612foundation.org/so-how-big-are-these-neos-anyway/		https://www.iau.org/public/themes/neo/nea3/
http://b612foundation.org/tunguska-meteorite-fragments-found/	b612foundation.org/tunguska-meteorite-fragments-found/	https://b612foundation.org/tunguska-stills/	https://www.iau.org/public/themes/neo/nea3/
http://becas.universiandecolombia.co/index.jsp	Universia Becas Colombia	http://becas.universiandecolombia.co/	http://www.poliagro.edu.co/portal/admisiones/ayudas-economicas.html
http://calgaryphil.com/event/arctic-ales-peer-gynt-northern-lights/	Arctic Tales: Peer Gynt & Northern Lights	http://calgaryphil.com/festival/	https://www.iau.org/public/publications/newsletter/2015_17/
http://decembermedia.com.au/	December Media	http://decembermedia.com.au/	https://www.eso.org/public/news/eso1329/
http://donaciones.uniandes.edu.co/s/1384/25/social.aspx?sid=1384&gid=25&pgid=789	Fondo de programas especiales (FOPRE)	https://catalogo.uniandes.edu.co/	https://uniandes.edu.co
http://eratosthenes.ea.gr/en/content/eratosthenes	Eratosthenes experiment	http://eratosthenes.ea.gr/en/content/experiment	https://www.iau.org/public/publications/newsletter/2015_13/
http://eratosthenes.ea.gr/en/content/eratosthenes-photo-contest	photo contest	http://eratosthenes2015.ea.gr/en/gallery	https://www.iau.org/public/publications/newsletter/2015_13/
http://events.rasc.ca	http://events.rasc.ca	https://www.rasc.ca/	https://www.iau.org/public/publications/newsletter/



			ter/2016_06/
http://gatheringthegathering.com/	http://gatheringthegathering.com/	https://findlaygathering.com/menus/the-gathering/	https://www.iau.org/public/publications/newsletter/2013_08/
http://geology.cwru.edu/~ansmet/	geology.cwru.edu/~ansmet/		https://www.iau.org/public/themes/neo/nea3/
http://gerencia.unal.edu.co/noticias-de-la-gerencia/article/novedades-del-proceso-contable-35.html?tx_ttnews%5BbackPid%5D=215&cHash=01dddc6ef0e8ad543314942bd61048dc		http://gerencia.unal.edu.co/noticias-de-la-gerencia/article/novedades-del-proceso-adquisicion-de-bienes-y-servicios-7.html?tx_ttnews%5Bpointer%5D=3&tx_ttnews%5BbackPid%5D=215&cHash=c888d886877c9fb4e21ba831c732bae3	http://gerencia.unal.edu.co/
http://gerhana.info/en	http://gerhana.info/en	https://gerhana.langitselatan.com/en/	https://www.iau.org/public/publications/newsletter/2016_01/
http://go.nasa.gov/1N9EP8u	http://go.nasa.gov/1N9EP8u	https://solarsystem.nasa.gov/educ/scientist-for-a-day	https://www.iau.org/public/publications/newsletter/2015_16/
http://iau.iteda.org	http://iau.iteda.org	http://web.archive.org/web/20130328104721/http://iau.iteda.org/	https://www.iau.org/scientific_meetings/iau_ga_2012/special_events/light_pollution/
http://iaucc1.frm.utn.edu.ar/wp-content/uploads/2012/11/IAUNL82.pdf	http://iaucc1.frm.utn.edu.ar/wp-content/uploads/2012/11/IAUNL82.pdf	http://iaucc1.frm.utn.edu.ar/?page_id=214	https://www.iau.org/public/publications/newsletter/2016_02/
http://image.scifest.net/2015/en/index.html	http://image.scifest.net/2015/en/index.html	http://web.archive.org/web/20150909172740/http://image.scifest.net/2015/en/index.html	https://www.iau.org/public/publications/newsletter/2015_13/
http://kiss.caltech.edu/study/asteroid/20120307_IEEE_Presentation.pdf	kiss.caltech.edu/study/asteroid/20120307_IEEE_Presentation.pdf	http://kiss.caltech.edu/papers/asteroid/papers/20120307_IEEE_Presentation.pdf	https://www.iau.org/public/themes/neo/nea3/
http://kiss.caltech.edu/study/asteroid/20120314_ESA_ESTEC.pdf	kiss.caltech.edu/study/asteroid/20120314_ESA_ESTEC.pdf	http://kiss.caltech.edu/papers/asteroid/papers/20120314_ESA_ESTEC.pdf	https://www.iau.org/public/themes/neo/nea3/
http://lasagraskysurvey.org/	lasagraskysurvey.org/	http://dig.do/lasagraskysurvey.org	https://www.iau.org/public/themes/neo/nea3/
http://macgillivrayfreemanfilms.com/	MacGillivray Freeman Films	http://macgillivrayfreeman.com/	https://www.eso.org/public/news/eso1329/

ENLACES EN PÁGINAS WEB



http://media.caltech.edu/press_releases/13524	media.caltech.edu/press_releases/13524		https://www.iau.org/public/themes/neo/nea3/
http://neo.jpl.nasa.gov/news/news126.html	neo.jpl.nasa.gov/news/news126.html		https://www.iau.org/public/themes/neo/nea2/
http://neo.jpl.nasa.gov/news/news170.html	neo.jpl.nasa.gov/news/news170.html		https://www.iau.org/public/themes/neo/nea3/
http://news.haverford.edu/blogs/unawe/	blog		http://es.unawe.org/about/team/
http://news.sciencemag.org/scienceinsider/2011/02/climate-science-asteroid-detection.html?ref=ra	news.sciencemag.org/scienceinsider/2011/02/climate-science-asteroid-detection.html?ref=ra	http://www.sciencemag.org/news/2011/02/climate-science-asteroid-detection-big-winners-nasa-budget	https://www.iau.org/public/themes/neo/nea3/
http://nightskylive.net/	Night Sky Live	http://slideplayer.com/slide/5364430/	https://www.eso.org/public/news/eso0648/
http://odeo.com/channel/386193/view/			https://www.spacetelescope.org/videos/archive/category/hubblecast/viwall/
http://pirlwww.lpl.arizona.edu/~guym/OSIRIS-REx.pdf	pirlwww.lpl.arizona.edu/~guym/OSIRIS-REx.pdf	https://www.lpl.arizona.edu/research/orex	https://www.iau.org/public/themes/neo/nea3/
http://sci.esa.int/jump.cfm?oid=48483	ESA website	http://sci.esa.int/xmm-newton/48483-an-old-galaxy-cluster-discovered-in-the-young-universe/	https://www.spacetelescope.org/announcements/ann1105/
http://shop.astronomerswithoutborders.org/collections/astronomers-without-borders/products/astronomers-without-borders-telescope	http://shop.astronomerswithoutborders.org/collections/astronomers-without-borders-telescope	http://shop.astronomerswithoutborders.org/collections/astronomers-without-borders-telescope	https://www.iau.org/public/publications/newsletter/2013_08/
http://space-awareness.ea.gr/en	Summer School	http://summerschools.ea.gr/content/space-awareness	http://www.unawe.org/updates/unawe-update-2016-02/
http://targetneo.jhuapl.edu/pdfs/2013_Presentations/Summary%20for%20SBAG%2010July2010.pdf	targetneo.jhuapl.edu/pdfs/2013_Presentations/Summary%20for%20SBAG%2010July2010.pdf	http://targetneo.jhuapl.edu/pdfs/2013_Presentations/Summary%20for%20SBAG%2010July2010.pdf	https://www.iau.org/public/themes/neo/nea3/
http://targetneo.jhuapl.edu/pdfs/sessions/TargetNEO-Session2-Harris.pdf	targetneo.jhuapl.edu/pdfs/sessions/TargetNEO-Session2-Harris.pdf	http://targetneo.jhuapl.edu/pdfs/2013_Presentations/session6/pdf/Session_2_Summary.pdf	https://www.iau.org/public/themes/neo/nea2/
http://targetneo.jhuapl.edu/pdfs/2013_Presentations/Summary%20for%20SBAG%2010July2010.pdf	targetneo.jhuapl.edu/pdfs/2013_Presentations/Summary%20for%20SBAG%2010July2010.pdf	http://targetneo.jhuapl.edu/pdfs/2013_Presentations/Summary%20for%20SBAG%2010July2010.pdf	https://www.iau.org/public/themes/neo/nea3/



edu/pdfs/sessions/TargetNEO-Session6-Jones.pdf	fs/sessions/TargetNEO-Session6-Jones.pdf	edu/pdfs/2013_Presentations/session6/pdf/Session_5_Summary.pdf	blic/themes/neo/nea3/
http://unisalud.unal.edu.co/index/sedes/medellin.html	Unisalud	http://unisalud.unal.edu.co/index/usuarios-medellin/horarios-de-atencion-medellin.html	http://medellin.unal.edu.co/
http://www.almaobservatory.org/en/press-room/press-releases/915-alma-reveals-planetary-construction-sites	ALMA	http://www.almaobservatory.org/en/press-release/alma-unveils-details-of-planet-formation-around-binary-star/	http://www.unawe.org/kids/unawe1604/
http://www.almaobservatory.org/en/visuals/live-from-the-alma-site/live-from-chajnantor	ALMA Array Operations Site (AOS)	http://www.almaobservatory.org/en/multimedia-en/live-from-chajnantor/	https://www.eso.org/public/announcements/ann14068/
http://www.almaobservatory.org/inauguration/	http://www.almaobservatory.org/inauguration/	http://www.almaobservatory.org/en/videos/scenes-from-the-alma-inauguration-day/	https://www.eso.org/public/announcements/ann13016/
http://www.astro.amu.edu.pl/~jopek/MDC2007/Roje/roje_lista.php?corobic_roje=0&sort_roje=0	www.astro.amu.edu.pl/~jopek/MDC2007/Roje/roje_lista.php?corobic_roje=0&sort_roje=0	http://pallas.astro.amu.edu.pl/~jopek/MDC2007/Roje/roje_lista.php?corobic_roje=0&sort_roje=0	https://www.iau.org/public/themes/neo/nea3/
http://www.astro.caltech.edu/~cosmos/	COSMOS Project Website	http://cosmos.astro.caltech.edu/	https://www.spacetelescope.org/news/heic0701/
http://www.astro.hr/spaceguard/	www.astro.hr/spaceguard/	http://www.astro.hr/vsa/2002/	https://www.iau.org/public/themes/neo/nea3/
http://www.astro.ubc.ca/MOST/	http://www.astro.ubc.ca/MOST/	http://www.astro.ubc.ca/	https://www.iau.org/science/meetings/future/loi_2011/loi05/
http://www.astro4dev.org/work-with-the-oad/opportunities-oad/	http://www.astro4dev.org/work-with-the-oad/opportunities-oad/	http://www.astro4dev.org/oad-visiting-fellowships/	https://www.iau.org/public/publications/newsletter/2014_06/
http://www.astroex.org/english/exercise3/	ESA/ESO Exercise	http://www.astroex.org/english/exercise3/pdf/5-Exercise3-high.pdf	https://www.spacetelescope.org/news/heic0414/
http://www.bu.edu/cas/2012/07/24/boston-university-joins-discovery-channel-telescope-partners-for-celebration-of-%E2%80%9Cfirst-	www.bu.edu/cas/2012/07/24/boston-university-joins-discovery-channel-telescope-partners-for-celebration-of-%E2%80%9Cfirst-	http://www.bu.edu/dct/	https://www.iau.org/public/themes/neo/nea3/



light%E2%80%9D/	light%E2%80%9D/		
http://www.capcolombia2016.org/	http://www.capcolombia2016.org/		https://www.iau.org/public/publications/newsletter/2015_17/
http://www.colorado.edu/news/features/ancient-asteroid-may-have-triggered-global-firestorm-earth	www.colorado.edu/news/features/ancient-asteroid-may-have-triggered-global-firestorm-earth	https://artsandsciences.colorado.edu/mag-old/2013/03/new-evidence-ancient-asteroid-caused-global-firestorm/	https://www.iau.org/public/themes/neo/nea3/
http://www.communicatingastronomy.org/cap2005/	CAP2005	https://www.communicatingastronomy.org/old/cap2005/	https://www.spacetelescope.org/about/history/sm3a_live_coverage/
http://www.communicatingastronomy.org/cap2013/index.html	CAP2013	https://www.communicatingastronomy.org/old/cap2013/index.html	https://www.spacetelescope.org/about/history/sm3a_live_coverage/
http://www.consejoderectores.cl/web/		http://www.consejoderectores.cl/	http://www.ascun.org.co/
http://www.crue.org/Paginas/Inicio.aspx		http://www.crue.org/SitePages/Inicio.aspx	http://www.ascun.org.co/
http://www.darksky.org/losingthedark	http://www.darksky.org/losingthedark	http://www.darksky.org/resources/losing-the-dark/	https://www.iau.org/public/publications/newsletter/2013_05/
http://www.e2v.com	http://www.e2v.com	https://www.e2v.com/	https://www.eso.org/public/news/eso1502/
http://www.eclipse-maps.com/Eclipse-Maps/Welcome_files/PS_E2011_November_25_Magnitude_and_greatest_eclipse.jpg	http://www.eclipse-maps.com/Eclipse-Maps/Welcome_files/PS_E2011_November_25_Magnitude_and_greatest_eclipse.jpg	http://www.eclipse-maps.com/Eclipse-Maps/Gallery/Pages/Partial_solar_eclipses_of_2011.html	https://www.iau.org/news/announcements/detail/ann11020/
http://www.esa.int/ssa/neo	www.esa.int/ssa/neo	http://www.esa.int/Our_Activities/Operations/Space_Situational_Awareness/Near-Earth_Objects_-_NEO_Segment	https://www.iau.org/public/themes/neo/nea3/
http://www.eso.org/e-elt/	E-ELT telescope		https://www.eso.org/public/announcements/ann14010/
http://www.eso.org/public/images/archive/search/?adv=&instrument=3&instrument=4	Other images taken with FORS	https://www.eso.org/public/images/archive/search/?adv=&instrument=3&instrument=4	https://www.eso.org/public/news/eso1707/
http://www.eso.org/public/news/archive/search/?adv=&facility=38	Link	https://www.eso.org/public/news/archive/search/?adv=&facility=65	https://www.eso.org/public/teles-instr/lasilla/rem/



http://www.eso.org/public/news/archive/search/?adv=&facility=40	Link	https://www.eso.org/public/news/archive/search/?adv=&facility=62	https://www.eso.org/public/teles-instr/lasilla/tarot/
http://www.eso.org/public/outreach/products/virtualtours/	ESO Virtual Tours	http://www.eso.org/public/outreach/	https://www.eso.org/public/announcements/ann14047/
http://www.eso.org/public/teles-instr/paranal/	Paranal Observatory	http://www.eso.org/public/teles-instr/paranal-observatory/vlt/	https://www.eso.org/public/news/eso1707/
http://www.eso.org/public/teles-instr/vlt/vlt-instr/fors/	FOcal Reducer and low dispersion Spectrograph 2	https://www.eso.org/public/teles-instr/paranal-observatory/vlt/vlt-instr/fors/	https://www.eso.org/public/news/eso1707/
http://www.eso.org/sci/observing/proposals/	http://www.eso.org/sci/observing/proposals/	https://www.eso.org/sci/observing/phase1.html	https://www.eso.org/public/announcements/ann11058/
http://www.fbbva.es/TLFU/tlfu/ing/microsites/premios/fronteras/galardonados/2011/ciencias.jsp	Press Release from the BBVA Foundation	https://www.frontiersofknowledgeawards-fbbva.es/version/edition_2011/	https://www.eso.org/public/announcements/ann12010/
http://www.floridatoday.com/article/20121130/COLUMNISTS0404/311300013/Stargazing-Asteroid-s-near-Earth-encounter-just-days-away	www.floridatoday.com/article/20121130/COLUMNISTS0404/311300013/Stargazing-Asteroid-s-near-Earth-encounter-just-days-away	http://web.archive.org/web/20150908022523/http://www.floridatoday.com/article/20121130/COLUMNISTS0404/311300013/Stargazing-Asteroid-s-near-Earth-encounter-just-days-away	https://www.iau.org/public/themes/neo/nea3/
http://www.forschung-garching.de/fileadmin/140915_KarteGarchingKomplett_RGB.pdf	map	http://www.forschung-garching.tum.de/fileadmin/w00btp/www/00_Startseite_normal/161015_KarteGarchingKomplett_RGB.pdf	https://www.eso.org/public/announcements/ann13070/
http://www.grandpublic.obspm.fr/2016-2023-Retour-sur-Terre-d	www.grandpublic.obspm.fr/2016-2023-Retour-sur-Terre-d	http://web.archive.org/web/20121012193413/http://www.grandpublic.obspm.fr/2016-2023-Retour-sur-Terre-d	https://www.iau.org/public/themes/neo/nea3/
http://www.harvard.edu/	Harvard University	http://www.harvard.edu/	https://www.eso.org/public/news/eso0323/
http://www.hofstra.edu/dcstars	www.hofstra.edu/dcstars	http://www.hofstra.edu/academics/colleges/hclas/physic/physic-nationalmall.html	https://www.iau.org/public/publications/newsletter/2016_07/



http://www.iafastro.com/index.html?title=Committee_on_Near_Earth_Objects	www.iafastro.com/index.html?title=Committee_on_Near_Earth_Objects		https://www.iau.org/public/themes/neo/nea3/
http://www.iau.org/grants_prizes/kavli_prize/	IAU. NASL and the Kavli Prize page	https://www.iau.org/science/grants_prizes/iau_grants/ga_events/ga_application_form/	https://www.iau.org/news/announcements/detail/ann12015/
http://www.iau.org/iau0702.486.0.html	Read more	https://www.iau.org/news/pressreleases/detail/iau0702/	https://www.iau.org/news/announcements/detail/ann07019/
http://www.iau.org/submissions/newcommissions/list/	http://www.iau.org/submissions/newcommissions/list/	https://www.iau.org/publications/e-newsletters/html/48/	https://www.iau.org/news/announcements/detail/ann15013/
http://www.icomos.org/studies	http://www.icomos.org/studies	http://www.icomos.org/studies/railways.pdf	https://www.iau.org/news/pressreleases/detail/iau1006/
http://www.icrar.org/research/the-universe-is-dying	Link to video fly-through of GAMA data	https://www.icrar.org/scientists-measure-slow-death-of-the-universe/	https://www.iau.org/news/pressreleases/detail/iau1509/
http://www.icsu.org/publications/about-icsu/icsu-brochure-august-2014/ICSU-brochure-2014.pdf	see PDF brochure	https://www.icsu.org/cms/2017/05/30-CSPR.pdf	https://www.iau.org/administration/international_org_relations/
http://www.ifsiroma.inaf.it/ewass2012/	http://www.ifsiroma.inaf.it/ewass2012/	https://web.archive.org/web/20120503221322/http://www.ifsiroma.inaf.it/ewass2012/	https://www.eso.org/public/announcements/ann12036/
http://www.insu.cnrs.fr/web/article/art.php?art=1618	press release		https://www.eso.org/public/news/eso0539/
http://www.ipa.nw.ru/PAGE/DEPFUND/LSBSS/engphem.htm	www.ipa.nw.ru/PAGE/DEPFUND/LSBSS/engphem.htm		https://www.iau.org/public/themes/neo/nea3/
http://www.iucaa.ernet.in/~scipop/atwst/participants.html	http://www.iucaa.ernet.in/~scipop/atwst/participants.html	http://web.archive.org/web/20140428120450/http://www.iucaa.ernet.in/~scipop/atwst/participants.html	https://www.iau.org/public/publications/newsletter/2014_01/
http://www.jpl.nasa.gov/asteroidwatch/newsfeatures.cfm?release=2011-053	www.jpl.nasa.gov/asteroidwatch/newsfeatures.cfm?release=2011-053	https://www.jpl.nasa.gov/news/news.php?release=2011-053	https://www.iau.org/public/themes/neo/nea3/
http://www.jpl.nasa.gov/asteroidwatch/newsfeatures.cfm?release=2011-	www.jpl.nasa.gov/asteroidwatch/newsfeatures.cfm?release=2011-296	https://www.jpl.nasa.gov/wise/newsfeatures.cfm?release=2011-296	https://www.iau.org/public/themes/neo/nea3/



296			
http://www.kakogawa-bunka.jp/movable-contents/2015/05/531.html	planetarium show in Kakogawa	http://www.kakogawa-bunka.jp/movable-contents/planetarium/	https://www.iau.org/public/publications/newsletter/2015_06/
http://www.khmw.nl/actueel/jong-talent-prijzen-2016/244	KHMW Announcement	https://khmw.nl/nieuwsarchief-232/	https://www.eso.org/public/announcements/ann16083/
http://www.kokuchpro.com/event/tanabata2015/	Gazing up into the Universe at Tanabata Night	http://web.archive.org/web/20150719233036/http://www.kokuchpro.com:80/event/tanabata2015	https://www.iau.org/public/publications/newsletter/2015_10/
http://www.lbl.gov/	Lawrence Berkley National Laboratory	http://www.lbl.gov/	https://www.eso.org/public/announcements/ann14084/
http://www.light2015.org/Home/Event-Programme/2015/Workshop/Romania---Educational-Astronomy.html	Educational Astronomy — A way towards learning societies	http://www.light2015.org/Home/Event-Programme/2015/Conference/Romania---Lighting-the-future-through-education.html	https://www.iau.org/public/publications/newsletter/2015_10/
http://www.loreal.com/_en/_ww/for-women-in-science.aspx	http://www.loreal.com/_en/_ww/for-women-in-science.aspx	http://www.loreal.my/	https://www.iau.org/news/pressreleases/detail/iau0907/
http://www.loreal.com/_en/_ww/newsMedia/news.aspx?NewsID=fa51cce3-98bd-4c52-8fa6-74097d2e9ba2	L'Oréal Press Release		https://www.iau.org/news/announcements/detail/ann10031/
http://www.lowell.edu/dct.php/	www.lowell.edu/dct.php/	http://ftp.lowell.edu/dct/index.php	https://www.iau.org/public/themes/neo/nea3/
http://www.lsst.org/lsst/news	www.lsst.org/lsst/news	https://www.lsst.org/news	https://www.iau.org/public/themes/neo/nea3/
http://www.lsst.org/overview/overview_v1.0.pdf	www.lsst.org/overview/overview_v1.0.pdf		https://www.iau.org/public/themes/neo/nea3/
http://www.meteorwatch.org/	http://www.meteorwatch.org/	http://www.meteorwatch.org/	https://www.eso.org/public/images/potw1033a/
http://www.mso.anu.edu.au/~rmn/	www.mso.anu.edu.au/~rmn/	http://rsaa.anu.edu.au/observatories/siding-spring-observatory	https://www.iau.org/public/themes/neo/nea3/
http://www.mtwilson.edu/michelson.php	Michelson prize		https://www.iau.org/news/announcements/detail/ann12017/
http://www.nature.com/doi/10.1038/natu	Scientific paper	http://www.nature.com/scitable/nated/topicpag	https://www.spacetelescope.org/news/heic041



re03006		e/scientific-papers-13815490	5/
http://www.noto.ira.inaf.it/index.php/dbbc-project	INAF/Noto Digital BaseBand Converter Project	http://web.archive.org/web/20160721062212/http://www.noto.ira.inaf.it:80/index.php/dbbc-project	https://www.eso.org/public/news/eso1229/
http://www.nwo.nl/nwo/home.nsf/pages/NWOP_5VNCW6	Spinoza Prize of 2.5 million euros	https://www.nwo.nl/en/news-and-events/news/2017/nwo-spinoza-prizes-awarded-to-eveline-crone-albert-heck-michel-orrit-and-alexander-van-oudenaarden.html	http://www.unawe.org/updates/unawe-update-1128/
http://www.orinoquia.unal.edu.co/bienestar.htm	Orinoquía	http://web.archive.org/web/20150320141302/http://www.orinoquia.unal.edu.co/bienestar.htm	http://www.bienestar.unal.edu.co/
http://www.pavconhecimento.pt/home/index.asp?acciao=changelang&lang=en	Pavilion of Knowledge - Ciência Viva	http://www.pavconhecimento.pt/home/?lang=2	https://www.spacetelescope.org/projects/20anniversary/events/
http://www.petermorse.com.au/	Peter Morse	http://www.petermorse.com.au/	https://supernova.eso.org/for-planetariums/
http://www.planetariomedellin.org/cap2016/english-version/registration/abstract-submission-and-grant-form/	application form	https://web.archive.org/web/20160313212247/http://www.planetariomedellin.org:80/cap2016/english-version/registration/abstract-submission-and-grant-form	https://www.iau.org/public/publications/newsletter/2015_19/
http://www.planetariomedellin.org/cap2016/english-version/registration/abstract-submission/	form	http://www.planetariomedellin.org/cap2016/english-version/programme/day-1/	https://www.iau.org/news/announcements/detail/ann15039/
http://www.planetariomedellin.org/cap2016/english-version/registration/registration-form/	form	http://www.planetariomedellin.org/cap2016/english-version/venues/medellin	https://www.iau.org/news/announcements/detail/ann15039/
http://www.planetariomedellin.org/shows/ fronteras-cuando-quieras-llegar-sera-momento-	http://www.planetariomedellin.org/shows/ fronteras-cuando-quieras-llegar-sera-momento-	http://www.planetariomedellin.org/shows-domo/fronteras	https://www.iau.org/public/publications/newsletter/2016_14/



de-partir/	de-partir/		
http://www.proyectomiradas.org/web/	Project Miradas	http://www.proyectomiradas.org/	http://www.unawe.org/updates/unawe-update-1411/
http://www.rgj.com/article/20120423/NEWS/304230032/Scientist-says-sound-signal-from-exploding-meteor- lasted-18-minutes?nclick_check=1	http://www.rgj.com/article/20120423/NEWS/304230032/Scientist-says-sound-signal-from-exploding-meteor- lasted-18-minutes?nclick_check=1		https://www.iau.org/public/themes/neo/nea3/
http://www.rmg.co.uk/about/international-activities/exhibitions	http://www.rmg.co.uk/about/international-activities/exhibitions	http://www.rmg.co.uk/work-services/what-we-do/international-touring-activities	https://www.iau.org/public/publications/newsletter/2014_02/
http://www.rmg.co.uk/whats-on/exhibitions/astronomy-photographer-of-the-year/exhibition	http://www.rmg.co.uk/whats-on/exhibitions/astronomy-photographer-of-the-year/exhibition	http://www.rmg.co.uk/royal-observatory/insight-astronomy-photographer-year/exhibition	https://www.iau.org/public/publications/newsletter/2014_02/
http://www.roe.ac.uk/ukatk/projects/upf/index.html		https://web.archive.org/web/20120802023716/http://www.roe.ac.uk/ukatk/projects/upf/index.html	https://www.iau.org/science/meetings/future/loi_2011/loi07/
http://www.ruinasdehuanchaca.cl/	Ruinas de Huanchaca	http://www.ruinasdehuanchaca.cl/	https://www.eso.org/public/announcements/ann14074/
http://www.scienceomega.com/article/1202/a-view-from-the-observatory-astronomic-collaboration-challenges	http://www.scienceomega.com/article/1202/a-view-from-the-observatory-astronomic-collaboration-challenges	http://web.archive.org/web/20130715112133/http://www.scienceomega.com:80/article/1202/a-view-from-the-observatory-astronomic-collaboration-challenges	https://www.iau.org/news/announcements/detail/ann13004/
http://www.servustv.com/cs/Satellite/About-ServusTV-empfangen/001259299587747?p=1259088496182	How to watch Servus TV	http://www.servustv.com/at/Service/TV-Empfang	https://www.eso.org/public/announcements/ann12088/
http://www.servustv.com/cs/Satellite/Article/Terra-Mater-011259461824540	Watch the documentary On-Demand at Servus TV	http://www.servustv.com/at/Videos	https://www.eso.org/public/announcements/ann12088/



http://www.silica.cl/	Sílica Networks Chile S.A.	http://www.silica.cl/	https://www.eso.org/public/announcements/ann15005/
http://www.space-x.eu/	Space Exploration Institute in Switzerland	http://www.space-x.eu/?space03_prod=get-eriacta-online	https://www.eso.org/public/announcements/ann13085/
http://www.stargazersclubwa.com.au/news/save-our-stars-free-stargazing-night	Save Our Stars		https://www.iau.org/public/publications/newsletter/2015_11/
http://www.stfc.ac.uk/astroprize	http://www.stfc.ac.uk/astroprize		https://www.iau.org/public/publications/newsletter/2014_0902/
http://www.telegraph.co.uk/science/space/8724019/Meteorite-soars-over-Peru.html	www.telegraph.co.uk/science/space/8724019/Meteorite-soars-over-Peru.html	http://www.telegraph.co.uk/news/science/space/8724019/Meteorite-soars-over-Peru.html	https://www.iau.org/public/themes/neo/nea3/
http://www.tes.co.uk/MyPublicProfile.aspx?uc=3600195&event=21&profileTab=resources	http://www.tes.co.uk/MyPublicProfile.aspx?uc=3600195&event=21&profileTab=resources		https://www.iau.org/public/publications/newsletter/2014_02/
http://www.toptica.com/	Toptica	http://www.toptica.com/	https://www.eso.org/public/announcements/ann16060/
http://www.tu-dresden.de/mwilr/space/rfs/Dokumente/Publikationen/NEO-MIPA-report.pdf	www.tu-dresden.de/mwilr/space/rfs/Dokumente/Publikationen/NEO-MIPA-report.pdf	https://web.archive.org/web/20110611084029/www.tu-dresden.de/mwilr/space/rfs/Dokumente/Publikationen/NEO-MIPA-report.pdf	https://www.iau.org/public/themes/neo/nea2/
http://www.uni-goettingen.de/de/presseinformation-europ%C3%A4isches-riesenteleskop-erh%C3%A4lt-modernsten-spektrografen-18032016/535417.html	Goettingen University		https://www.eso.org/public/announcements/ann16017/
http://www.universetoday.com/11734/never-a-star-did-supermassive-black-holes-form-directly/	direct-collapse black hole		https://www.eso.org/public/news/eso1524/
http://www.universetoday.com/81912/upcomin	http://www.universetoday.com/81912/upcomin		https://www.iau.org/news/announcements/det



http://www.windowsphone.com/en-us/store/app/european-southern-observatory/6a9d5ab4-132f-4483-9581-6ba5ed8503ce	Windows Phone 8		ail/ann10037/
http://www.wuala.com/IAAbackup/Big%20Files%20Flagstaff/?key=G6FttSjkXG9T	www.wuala.com/IAAbackup/Big%20Files%20Flagstaff/?key=G6FttSjkXG9T		https://www.iau.org/public/themes/neo/nea3/
https://impactotic.co/blog2__trashed/	Curaduría: lo más relevante de la Web		https://impactotic.co/
https://www.darkenergy-survey.org/	Dark Energy Survey	https://www.darkenergy-survey.org/	https://www.eso.org/public/announcements/ann16055/
https://www.eso.org/public/outreach/gems/	ESO's Cosmic Gems programme	https://www.eso.org/public/outreach/gems/	https://www.eso.org/public/news/eso1707/
https://www.eso.org/public/teles-instr/vlt/vlt-names/	Antu	https://www.eso.org/public/teles-instr/paranal-observatory/vlt/vlt-names/	https://www.eso.org/public/news/eso1707/
https://www.iau.org/science/scientific_bodies/commissions/14/		https://www.iau.org/science/scientific_bodies/commissions/14/members/	https://www.iau.org/science/scientific_bodies/divisions/F/info/commissions/
https://www.iau.org/science/scientific_bodies/commissions/19/	Commission 19 / A2 Rotation of the Earth	https://www.iau.org/science/scientific_bodies/past_commissions/19/	https://www.iau.org/science/scientific_bodies/divisions/A/info/news/
https://www.iau.org/science/scientific_bodies/commissions/31/	Commission 31 Time	https://www.iau.org/science/scientific_bodies/past_commissions/31/	https://www.iau.org/science/scientific_bodies/divisions/A/info/news/
https://www.iau.org/science/scientific_bodies/commissions/8/	Commission 8 / A1 Astrometry		https://www.iau.org/science/scientific_bodies/divisions/A/info/news/
https://www.iau.org/science/scientific_bodies/working_groups/211/	IAU Webpage		https://www.iau.org/science/scientific_bodies/divisions/F/info/wgs/
https://www.iau.org/science/scientific_bodies/working_groups/214/	WG Nominal Units for Stellar & Planetary Astronomy		https://www.iau.org/science/scientific_bodies/divisions/A/info/news/
http://adsabs.harvard.edu/cgi-bin/basic_connect?qsearch=Ara+OB1&version=1	Research papers on Ara OB1	http://adsabs.harvard.edu/cgi-bin/basic_connect?qsearch=Ara+OB1&version=1	https://www.eso.org/public/news/eso1510/



http://almaobservatory.org/en/press-room/announcements-events/427-alma-telescope-upgrade-to-power-new-science	Announcement from the Joint ALMA Observatory	http://almaobservatory.org/en/press-room/announcements-events/427-alma-telescope-upgrade-to-power-new-science	https://www.eso.org/public/announcements/ann12042/
http://almaobservatory.org/en/press-room/announcements-events/884-grand-exhibit-on-alma-opens-in-santiago	Announcement on the JAO ALMA site		https://www.eso.org/public/announcements/ann15077/
http://articles.adsabs.harvard.edu/cgi-bin/nph-iarticle_query?1966ApJ..144..259A&data_type=PDF_HIGH&whole_paper=YES&type=PRINTER&filetype=.pdf	Abell Catalogue of Planetary Nebulae	https://ui.adsabs.harvard.edu/	https://www.eso.org/public/news/eso1412/
http://astronomy2015.org/press_tour	Press Tour Information		https://www.iau.org/news/announcements/detail/ann15019/
http://astronomy2015.org/press_tour	here		https://www.iau.org/news/announcements/detail/ann15019/
http://atacamaphoto.com	atacamaphoto.com	http://atacamaphoto.com/sitemap1.htm	https://www.eso.org/public/announcements/ann15032/
http://b612foundation.org/downloads/pressRelease_6_28.zip	b612foundation.org/downloads/pressRelease_6_28.zip	https://b612foundation.org/wp-content/uploads/2016/05/B612-Foundation-Sentinel-Space-Telescope.pdf	https://www.iau.org/public/themes/neo/nea3/
http://b612foundation.org/media/sentinelmission/	b612foundation.org/media/sentinelmission/	https://b612foundation.org/sentinel/	https://www.iau.org/public/themes/neo/nea3/
http://b612foundation.org/rustys-post-about-keyholes/	b612foundation.org/rustys-post-about-keyholes/	https://b612foundation.org/asteroid-day-recap-webinar-with-rusty-schweickart/	https://www.iau.org/public/themes/neo/nea3/
http://b612foundation.org/rustys-post-an-asteroid-with-our-address/	b612foundation.org/rustys-post-an-asteroid-with-our-address/	https://b612foundation.org/asteroid-day-recap-webinar-with-rusty-schweickart/	https://www.iau.org/public/themes/neo/nea3/
http://b612foundation.org/rustys-post-an-asteroid-with-our-address/	b612foundation.org/rustys-post-an-asteroid-with-our-address/	https://b612foundation.org/asteroid-day-recap-webinar-with-rusty-schweickart/	https://www.iau.org/public/themes/neo/nea3/
http://b612foundation.org/rustys-post-an-asteroid-with-our-address/	b612foundation.org/rustys-post-an-asteroid-with-our-address/	https://b612foundation.org/asteroid-day-recap-webinar-with-rusty-schweickart/	https://www.iau.org/public/themes/neo/nea3/



rg/rustys-post-changing-the-orbit-of-an-asteroid/	ys-post-changing-the-orbit-of-an-asteroid/	org/scientists-worldwide-mark-first-annual-asteroid-day-with-stark-warnings/	b612foundation.org/public/themes/neo/nea3/
http://b612foundation.org/tag/planetary-defense-conference/	b612foundation.org/tag/planetary-defense-conference/	https://b612foundation.org/b612s-progress-report-vol-1-engaging-the-world-in-planetary-defense/	https://www.iau.org/public/themes/neo/nea3/
http://b612foundation.org/tedxmarinedsept2012/	b612foundation.org/tedxmarinedsept2012/	https://b612foundation.org/tedxmarin-talk-with-dr-ed-lu-changing-the-course-of-the-solar-system/	https://www.iau.org/public/themes/neo/nea3/
http://capcolombia2016.org/	Communicating Astronomy with the Public 2016 website		https://www.iau.org/news/announcements/detail/ann15001/
http://cosmiclog.msnbc.msn.com/_news/2011/11/16/8845471-asteroid-debate-rises-to-next-level	cosmiclog.msnbc.msn.com/_news/2011/11/16/8845471-asteroid-debate-rises-to-next-level		https://www.iau.org/public/themes/neo/nea3/
http://creativecommons.org/licenses/by/3.0/	Creative Commons Attribution 3.0 Unported license	https://creativecommons.org/licenses/by/3.0/	https://www.iau.org/public/themes/constellations/
http://csp.obs.carnegiescience.edu/data/lowzSNe/SN2004dt	SN2004dt	http://csp.obs.carnegiescience.edu/data/lowzSNe/SN2004dt/SN04dt_b01_DUP_WF_19sep04.dat	https://www.eso.org/public/images/potw1332a/
http://decembermedia.com.au/2012/07/december-media-starts-production-on-the-big-one-as-big-as-it-gets/	December Media	http://decembermedia.com.au/2012/07/december-media-starts-production-on-the-big-one-as-big-as-it-gets/	https://www.eso.org/public/announcements/ann12089/
http://emotimo.com/	eMotimo	http://emotimo.com/	https://www.eso.org/public/images/potw1414a/
http://flandrau.org/exhibits/iaaa	The Heritage of Astronomical Art in Arizona	http://iaaatucson.org/the-heritage-of-astronomical-art-in-arizona/	https://www.iau.org/public/publications/newsletter/2015_08/
http://galileo-mobile.org/	http://galileo-mobile.org/		https://www.iau.org/public/publications/newsletter/2014_05/
http://geology.missouri.edu/resources/rockid.html	University of Missouri's Geological Department guidance		https://www.iau.org/public/themes/discoveries/
http://gerhana.langitsel	Chasing the Shadows	http://gerhana.langitsel	https://www.iau.org/pu



atan.com/en/photo-exhibitions-chasing-the-shadows/		atan.com/en/photoexhibitions/	blic/publications/newsletter/2016_08/
http://hubbledev.stsci.edu/newscenter/archive/releases/2015/23/	Link to hubblesite release	http://hubble.stsci.edu/search/	https://www.spacetelescope.org/news/heic1513/
http://hubbledev.stsci.edu/newscenter/archive/releases/2016/08/	Link to Hubblesite release	http://hubble.stsci.edu/search/	https://www.spacetelescope.org/images/potw1611a/
http://indico.universe-cluster.de/indico/conferenceDisplay.py?confId=3274	A flyer with the programme. travel information and further details is available	http://indico.universe-cluster.de/indico/confLogin.py?confId=3274	https://www.eso.org/public/announcements/ann15008/
http://kiss.caltech.edu/workshops/asteroid2011/index.html	kiss.caltech.edu/workshops/asteroid2011/index.html	http://web.archive.org/web/20170610044617/http://kiss.caltech.edu/workshops/asteroid2011/index.html	https://www.iau.org/public/themes/neo/nea3/
http://michellebachelet.cl	Incoming President Bachelet web pages	http://michellebachelet.cl/	https://www.eso.org/public/announcements/ann14017/
http://monuauniandes.org	MONUA	http://monuauniandes.org/united-nations/	https://uniandes.edu.co
http://most.astro.ubc.ca/	MOST	http://most.astro.ubc.ca/news28.html	https://www.eso.org/public/news/eso0215/
http://muse-vlt.eu/blog/MUSE-Comm/Blog/Blog.html	followed here	http://muse-vlt.eu/blog/MUSE-Comm/Blog/Archive.html	https://www.eso.org/public/news/eso1407/
http://neo.jpl.nasa.gov/news/news131.html	neo.jpl.nasa.gov/news/news131.html		https://www.iau.org/public/themes/neo/nea3/
http://neo.jpl.nasa.gov/news/news171.html	neo.jpl.nasa.gov/news/news171.html		https://www.iau.org/public/themes/neo/nea3/

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



APÉNDICE D. Registro de Software.

		MINISTERIO DEL INTERIOR DIRECCION NACIONAL DE DERECHO DE AUTOR UNIDAD ADMINISTRATIVA ESPECIAL OFICINA DE REGISTRO		Libro - Tomo - Partida 13-62-29
CERTIFICADO DE REGISTRO DE SOPORTE LOGICO - SOFTWARE				Fecha Registro 23-jul-2017
Página 1 de 2				
1. DATOS DE LAS PERSONAS				
AUTOR				
Nombres y Apellidos	JAVIER MAURICIO ENCISO MORENO	No de identificación CC	80097702	
Nacional de	COLOMBIA			
Dirección	CR 9 37 04 CS 61 PARQUES DE SEVII	Ciudad:	VILLAVICENCIO	
AUTOR				
Nombres y Apellidos	MAXIMILIANO CARDENAS GELVES	No de identificación CC	1121925871	
Nacional de	COLOMBIA	F. De Defunción		
	LOTE 58 VRDA. BARCELONA, PARCEI	Ciudad:	VILLAVICENCIO	
AUTOR				
Nombres y Apellidos	EMIRALDO ANDRÉS LOZANO SÁNCHEZ	No de identificación CC	1121930903	
Nacional de	COLOMBIA			
Dirección	--	Ciudad:	VILLAVICENCIO	
AUTOR				
Nombres y Apellidos	ALFONSO ENRIQUE DÍAZ MALAVER	No de identificación CC	1049605593	
Nacional de	COLOMBIA	F. De Defunción		
	CALLE 3 SUR #28A-46 MZ 24 CS 12B	Ciudad:	VILLAVICENCIO	
AUTOR				
Nombres y Apellidos	MARIA ALEJANDRA ROJAS RIOS	No de identificación CC	40342362	
Nacional de	COLOMBIA			
Dirección	--	Ciudad:	VILLAVICENCIO	
TITULAR DERECHO PATRIMONIAL				
Nombres y Apellidos	JAVIER MAURICIO ENCISO MORENO	No de identificación CC	80097702	
Nacional de	COLOMBIA			
Dirección	CR 9 37 04 CS 61 PARQUES DE SEVII	Ciudad:	VILLAVICENCIO	
PRODUCTOR				
Nombres y Apellidos	JAVIER MAURICIO ENCISO MORENO	No de identificación CC	80097702	
Nacional de	COLOMBIA			
Dirección	CR 9 37 04 CS 61 PARQUES DE SEVII	Ciudad:	VILLAVICENCIO	
2. DATOS DE LA OBRA				
Título Original	BYE BYE 404			
Año de Creación	2017	País de Origen	COLOMBIA	Año Edición
CLASE DE OBRA	INEDITA			
CARACTER DE LA OBRA	OBRA ORIGINARIA			
CARACTER DE LA OBRA	OBRA COLECTIVA			
ELEMENTOS APORTADOS DE SOPORTE LOGICO	PROGRAMA DE COMPUTADOR			
3. DESCRIPCIÓN DE LA OBRA APLICACIÓN WEB PARA DETECTAR Y REPARAR HIPERVÍNCULOS ROTOS EN PÁGINAS WEB.				
4. OBSERVACIONES GENERALES DE LA OBRA				

ALGORITMO DE BÚSQUEDA Y RECOMENDACIÓN PARA REPARAR ENLACES EN PÁGINAS WEB



	MINISTERIO DEL INTERIOR DIRECCION NACIONAL DE DERECHO DE AUTOR UNIDAD ADMINISTRATIVA ESPECIAL OFICINA DE REGISTRO		Libro - Tomo - Partida 13-62-29
	CERTIFICADO DE REGISTRO DE SOPORTE LOGICO - SOFTWARE		Fecha Registro 23-jul-2017
Página 2 de 2			
5. DATOS DEL SOLICITANTE			
Nombres y Apellidos	JAVIER MAURICIO ENCISO MORENO	No de Identificación	80097702
Nacional de	COLOMBIA	Medio Radicación	REGISTRO EN LINEA
Dirección	CR 9 37 04 CS 61 PARQUES DE SEVILLA 1	Ciudad	VILLAVICENCIO
Correo electrónico	J4R.E4O@GMAIL.COM	Teléfono	314 487 8000
En representación de	EN NOMBRE PROPIO	Radicación de entrada	1-2017-55958
 MANUEL ANTONIO MORA CUELLAR JEFE OFICINA DE REGISTRO			
MZP			

Nota: El derecho de autor protege exclusivamente la forma mediante la cual las ideas del autor son descritas, explicadas, ilustradas o incorporadas a las obras. No son objeto de protección las ideas contenidas en las obras literarias y artísticas, o el contenido ideológico o técnico de las obras científicas, ni su aprovechamiento industrial o comercial (artículo 7o. de la Decisión 351 de 1993).

APÉNDICE E. Presentación en Villadevs

Villadevs es un foro abierto para la discusión de los últimos desarrollos tecnológicos en la región del Meta y Llanos Orientales. Está patrocinado en el contexto del plan nacional vive digital por ParqueSoft Meta, el Ministerio Colombiano de TIC y la oficina mayor de Villavicencio.



Ilustración 8. Ing. M.Sc. Javier Enciso introduciendo la Aplicación BYE BYE 404. Fuente: Villadevs.

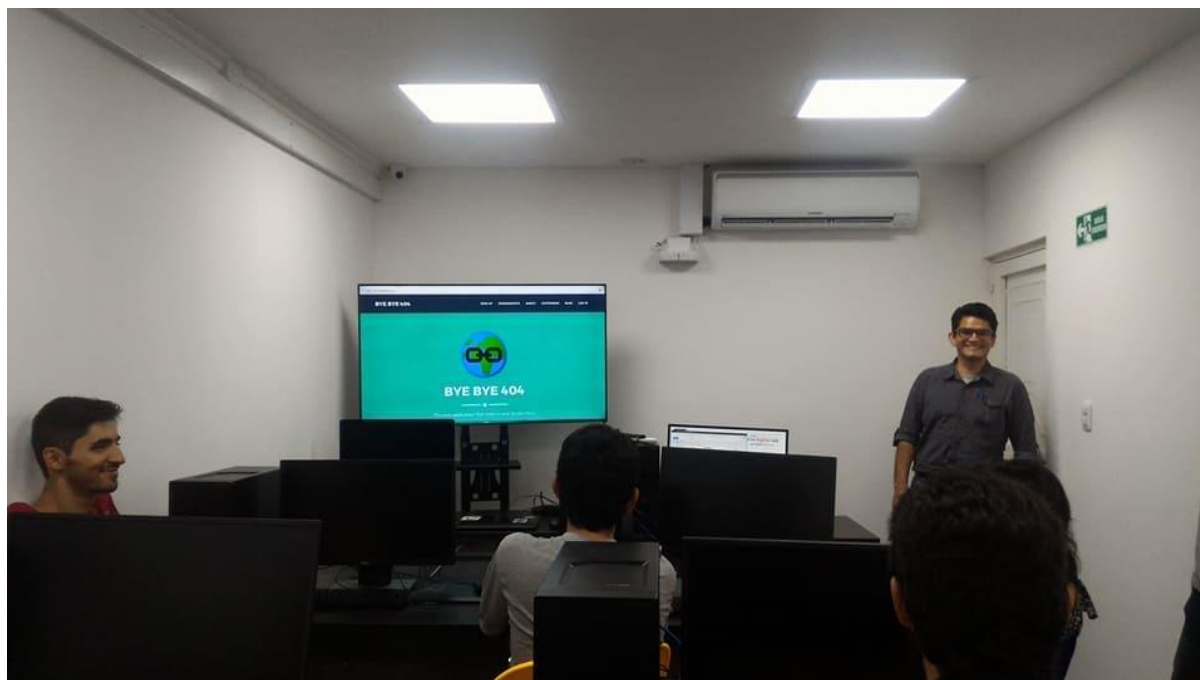


Ilustración 9. Pasante Maximiliano Cárdenas presentando forma de uso de BYE BYE 404. Fuente: Villadevs.



Ilustración 10. Pasante Emiraldo Lozano presentando algoritmo de recomendaciones implementado en BYE BYE 404. Fuente: Villadevs.



APÉNDICE F. Mención a BYE BYE 404.

IAU es un cliente de BYE BYE 404 que recomienda el uso de la aplicación a la comunidad internacional.

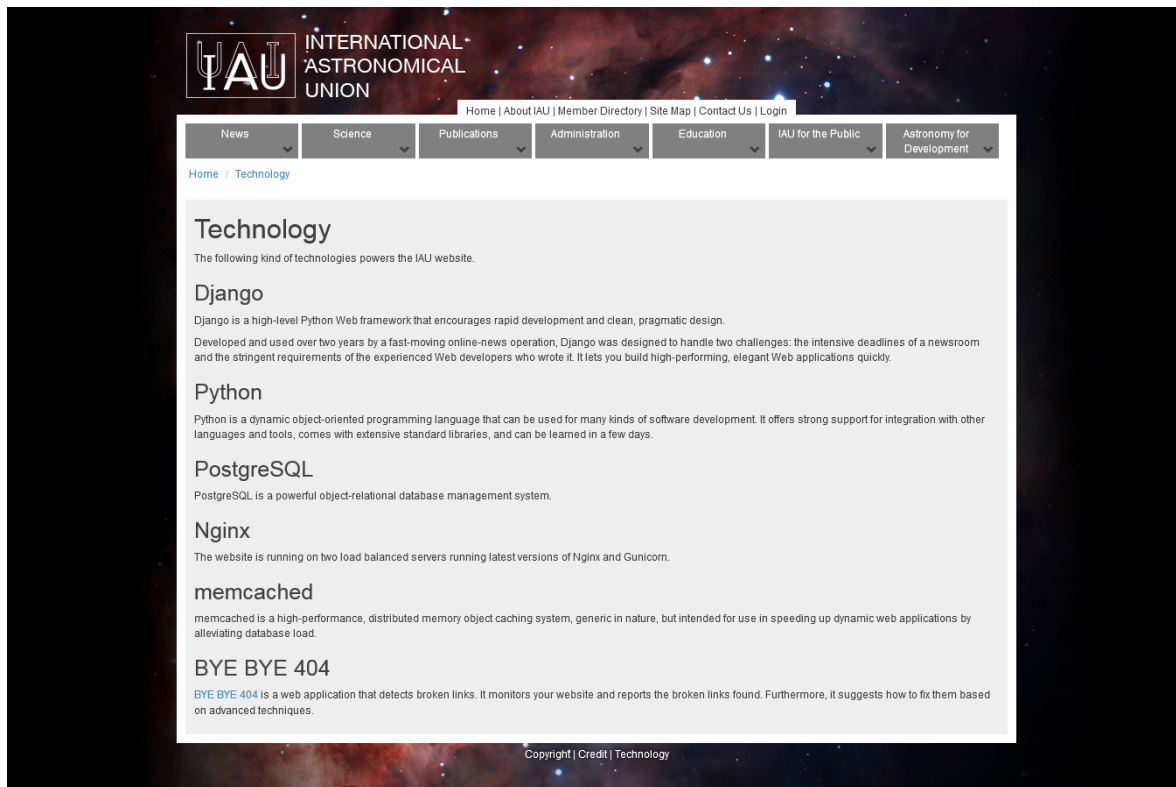


Ilustración 11. Recomendación de IAU. Fuente: IAU.