



UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA
La Universidad Católica de Loja

ÁREA TÉCNICA

**TÍTULO DE INGENIERO EN SISTEMAS INFORMÁTICOS Y
COMPUTACIÓN**

Plug-in para la plataforma OJS

TRABAJO DE TITULACIÓN

AUTOR: Gutiérrez Roa, Rommel Augusto

DIRECTOR: Torres Díaz, Juan Carlos, Ing.

LOJA – ECUADOR

2016



Esta versión digital, ha sido acreditada bajo la licencia Creative Commons 4.0, CC BY-NY-SA: Reconocimiento-No comercial-Compartir igual; la cual permite copiar, distribuir y comunicar públicamente la obra, mientras se reconozca la autoría original, no se utilice con fines comerciales y se permiten obras derivadas, siempre que mantenga la misma licencia al ser divulgada. <http://creativecommons.org/licenses/by-nc-sa/4.0/deed.es>

Septiembre, 2016

APROBACIÓN DEL DIRECTOR DEL TRABAJO DE TITULACIÓN

Ingeniero.

Juan Carlos Torres Díaz.

DOCENTE DE LA TITULACIÓN

De mi consideración:

El presente trabajo de titulación: "**Plug-in para la plataforma OJS**", realizado por **Rommel Augusto Gutiérrez Roa**; ha sido orientado y revisado durante su ejecución, por cuanto se aprueba la presentación del mismo.

Loja, abril de 2016

f).....

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS

“Yo **Rommel Augusto Gutiérrez Roa** declaro ser autor del presente trabajo de titulación: “Plug-in para la plataforma OJS”, de la Titulación Ingeniería en Sistemas y Computación, siendo Juan Carlos Torres Díaz director del presente trabajo; y eximo expresamente a la Universidad Técnica Particular de Loja y a sus representantes legales de posibles reclamos o acciones legales. Además certifico que las ideas, conceptos, procedimientos y resultados vertidos en el presente trabajo investigativo, son de mi exclusiva responsabilidad.

Adicionalmente declaro conocer y aceptar la disposición del Art. 88 del Estatuto Orgánico de la Universidad Técnica Particular de Loja que en su parte pertinente textualmente dice: “Forman parte del patrimonio de la Universidad la propiedad intelectual de investigaciones, trabajos científicos o técnicos y tesis de grado o trabajos de titulación que se realicen con el apoyo financiero, académico o institucional (operativo) de la Universidad”

f).....

Autor: Rommel Augusto Gutiérrez Roa

Cédula: **1104338908**

DEDICATORIA

Dedico este trabajo:

A las mentes anhelantes del conocimiento colaborativo.

A aquellas personas que no se conforman únicamente con vivir de teorías escritas en papel, sino que las ponen en práctica para comprobar su resultado.

A los auto didactas, que aprenden gracias a su propio interés y comparten el conocimiento adquirido.

AGRADECIMIENTO

A mi madre, a mi padre y a mi hermana quienes me apoyaron durante todo mi camino académico, y aún lo siguen haciendo en el camino de la vida.

A demás familiares y amigos que de alguna manera supieron llegar con sus palabras de aliento.

A la Universidad Técnica Particular de Loja que mediante sus profesores supieron impartir conocimiento. En especial al Ing. Juan Carlos Torres Díaz por su guía, aliento, dedicación y paciencia

Rommel Augusto Gutiérrez Roa

ÍNDICE

APROBACIÓN DEL DIRECTOR DEL TRABAJO DE TITULACIÓN	ii
DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS	iii
DEDICATORIA.....	iv
AGRADECIMIENTO.....	v
ÍNDICE	vi
ÍNDICE DE IMÁGENES	ix
ÍNDICE DE TABLAS	xii
RESUMEN	1
ABSTRACT	2
INTRODUCCIÓN.	3
CAPITULO I: MARCO TEÓRICO.....	1
1. Conceptos esenciales.....	1
1.1. Administración de contenidos (CM).....	1
1.1.1. CMS: Content Management System.....	2
1.1.2. WCMS: Web Content Management System.	2
1.2. Gestión de la información.	3
1.2.1. Open Access.....	4
1.2.2. Public Knowledge Project.....	4
1.2.3. OJS	5
1.2.3.1. Uso de OJS.....	5
1.2.3.2. Casos de éxito implementando OJS en Sudamérica.	8
1.2.4. E –Marketing.....	9
1.2.4.1. E-mail Marketing.	9
1.2.4.2. Search Engine Optimization.....	10
1.2.4.3. Article Level Metrics	10
1.2.5. Web Scraping.....	11
1.2.6. DOI.....	11
1.2.6.1. Sintaxis de nombramiento DOI.	12
1.2.6.2. Ventajas del sistema DOI.....	12
1.3. Metodologías de Desarrollo.	13
1.3.1. Metodologías Ágiles.....	13
1.3.2. Metodología XP.....	15
CAPITULO II: ESTADO DEL ARTE	16
2. Trabajos relacionados sobre enriquecimiento de información.....	17
2.1. Altmetric: enriching scholarly content with article-level discussion and metrics.....	17

2.2.	A graphical article-level metric for intuitive comparison of large-scale literatures	19
2.3.	Métodos y técnicas para el cálculo y visualización de métricas de historiales en artículos de wikis.....	20
CAPITULO III: PROBLEMÁTICA Y SOLUCIÓN PROPUESTA		22
3.	Problemática y Objetivos.....	23
3.1.	Problemática actual.....	23
3.1.1.	Propuesta para OJS.....	24
3.1.2.	Propuesta para marketing.....	25
3.2.	Objetivos.....	27
CAPITULO IV: DESARROLLO E IMPLEMENTACIÓN		28
4.	Desarrollo e Implementación.....	29
4.1.	Propósito.....	29
4.2.	Metodología.....	30
4.3.	Desarrollo del plugin.....	31
4.3.1.	Planificación.....	31
4.3.1.1.	Historias de usuario	31
4.3.2.	Diseño.....	33
4.3.2.1.	Versiones e Iteraciones.....	35
4.3.3.	Desarrollo.....	38
4.3.3.1.	Extracción de Métricas.....	38
4.3.3.2.	ELSEVIER.....	38
4.3.3.2.1.	Altmetric.....	40
4.3.3.3.	Procesamiento de datos.....	42
4.3.3.4.	Pantallas y visualización.....	44
4.3.4.	Pruebas.....	49
4.4.	Desarrollo de la Herramienta de Marketing.....	50
4.4.1.	Planificación.....	51
4.4.1.1.	Historias de usuario	51
4.4.2.	Diseño.....	52
4.4.2.1.	Versiones e Iteraciones.....	52
4.4.3.	Desarrollo.....	55
4.4.3.1.	Análisis de la estructura HTML del resumen del artículo.....	56
4.4.3.2.	Pantallas y visualización.....	58
4.4.3.3.	Procesamiento de datos.....	59
4.4.4.	Pruebas.....	60
4.4.4.1.	Pruebas Funcionales.....	60

CONCLUSIONES.....	65
RECOMENDACIONES	66
Bibliografía	67
ANEXOS	69
A. Configuraciones iniciales.....	70
B. Ejemplo de archivo CSV que se puede exportar con la herramienta de marketing.....	79
C. Manual de instalación de la herramienta de marketing.....	80
D. Plugin. Manual de Programador	82
E. Plugin. Manual de Usuario.....	94
F. Marketing. Manual de Programador.	99
G. Marketing. Manual de Usuario.....	102

ÍNDICE DE IMÁGENES

Figura 1: Localización de Revistas usando OJS en el año 1990.....	6
Figura 2: Localización de Revistas usando OJS en el año 2015.....	6
Figura 3: Gráfico que representa el crecimiento del uso de OJS desde 1990 hasta el año actual.....	7
Figura 4: Gráfico que representa el crecimiento del número de artículos publicados mediante OJS desde 1990.	7
Figura 5: Gráfico que representa el crecimiento del número de instalaciones desde 1990 que hacen uso de OJS para sus publicaciones.....	8
Figura 6: Estructura del DOI.....	12
Figura 7: Volumen de actividad de servicios en línea que involucran un enlace a un artículo académico, rastreado por Altmetric	17
Figura 8: detalles de Altmetric para el paper ENCODE.....	18
Figura 9: Gráfica del incremento de uso de Altmetric.....	19
Figura 10: Papers de PLoS Biology publicados durante 2003 y 2014 por SNM	20
Figura 11: Métrica de revisiones por usuario.....	20
Figura 12: visualización de un artículo con la actual visualización en OJS.	24
Figura 13: presentación de un artículo en Plos Biology	34
Figura 14: Barra de Navegación de OJS.....	38
Figura 15: URL formada para la consulta del DOI en el API de Altmetric.	39
Figura 16: XML de respuesta de SCOPUS a través del navegador Firefox.	39
Figura 17: Código para obtener información de ELSEVIER.....	40
Figura 18: Información del DOI mediante el API Altmetric.	41
Figura 19: URL formada para la consulta del DOI en el API de Altmetric.	41
Figura 20: JSON de Altmetric a través de http://jsonviewer.stack.hu/	42
Figura 21: (izq.) Tabla y gráfica de las métricas de OJS. (der.) Recuadro de información externa.....	43
Figura 22: Artículo de la revista de ejemplo. Pestaña principal Abstract.	44
Figura 23: Pestaña métricas de la revista de ejemplo.....	45
Figura 24: Pestaña autores de la revista de ejemplo.	46
Figura 25: Pestaña comentarios de la revista de ejemplo.....	47
Figura 26: Ventana emergente para elegir la forma de citar el artículo.	47
Figura 27: Acción de copia de la cita seleccionada.....	48
Figura 28: Opción de compartir artículo.	48
Figura 29: Ejemplo del resumen de un artículo	56
Figura 30: Ejemplo HTML del resumen de un artículo	56
Figura 31: Referencias de un artículo en SCOPUS	57
Figura 32: ventana de ingreso a la herramienta de marketing.	58
Figura 33: formas de ingresar DOI para procesar.....	58
Figura 34: Ejemplo del archivo con lista de DOI.....	59
Figura 35: XML de respuesta.	59
Figura 36: Fin de la ejecución de la herramienta de marketing.....	60
Figura 37: Instalación de PHP en Synaptic.	70
Figura 38: Comando de instalación para PHP5 en la consola (terminal).	71
Figura 39: Instalación de MySQL a través de Synaptic.....	71
Figura 40: Comandos de instalación de MySQL a través del Terminal.....	72
Figura 41: Petición de contraseña para MySQL Server 5.5.	72
Figura 42: Confirmación de la contraseña introducida para MySQL Server 5.5.....	73
Figura 43: Creación de la base de datos para OJS.....	73

Figura 44: Asignación de permisos a un usuario para la base de datos de OJS.	74
Figura 45: Comando de instalación de apache.	74
Figura 46: Captura del mensaje que demuestra que apache2 está instalado.	74
Figura 47: Archivo de verificación de php	75
Figura 48: información de PHP al ingresar en el navegador al archivo de información.	75
Figura 49: Página de PKP para la descarga de OJS.	76
Figura 50: Descompresión del archivo que contiene OJS.	76
Figura 51: comandos para creación de carpetas y permisos de archivos.	77
Figura 52: Copia de archivos PJS a la carpeta de Apache y cambio de usuario de archivos.	77
Figura 53: Página de instalación de OJS.	78
Figura 54: Ejemplo de archivo CSV obtenido con la herramienta de marketing.	79
Figura 55: Carpetas y archivos de OJS.	82
Figura 56: Archivos básicos dentro de la carpeta "utplMetrics".	83
Figura 57: archivo "index.php"	84
Figura 58: archivo "version.xml"	84
Figura 59: archivo "utplPlugin.inc.php".	85
Figura 60: Hook "Template Manager::display".	86
Figura 61: formulario para crear una cuenta de ELSEVIER Developers.	87
Figura 62: "My Projects", sección para obtener la API Key.	87
Figura 63: Documentación en línea del API para la búsqueda en SCOPUS.	88
Figura 64: URL formada para la consulta del DOI en el API de Altmetric.	88
Figura 65: Ejemplo XML de respuesta de SCOPUS a través del navegador Firefox.	89
Figura 66: Código para obtener información de ELSEVIER.	90
Figura 67: Obtención de la API Key de Altmetric.	91
Figura 68: Información del DOI mediante el API Altmetric.	91
Figura 69: Documentación del API de Altmetric.	92
Figura 70: URL formada para la consulta del DOI en el API de Altmetric.	92
Figura 71: JSON de Altmetric.	93
Figura 72: Ingreso al Administrador de la revista	94
Figura 73: Páginas de administración.	95
Figura 74: Administración de plugins.	95
Figura 75: Formulario de instalación de plugin.	96
Figura 76: Mensaje de éxito al instalar plugin.	96
Figura 77: Administración de plugins, Plugins genéricos	97
Figura 78: Habilitación de plugin utplMetrics.	97
Figura 79: Confirmación de habilitado.	98
Figura 80: Configuración de utplMetrics.	98
Figura 81: Notificación de configuración cambiada.	98
Figura 82: Documentación, obtención del resumen del artículo mediante el API de ELSEVIER.	99
Figura 83: Ejemplo del resumen de un artículo	100
Figura 84: Ejemplo HTML del resumen de un artículo	100
Figura 85: Referencias de un artículo en SCOPUS	101
Figura 86: ventana de ingreso a la herramienta de marketing.	102
Figura 87: formas de ingresar DOI para procesar.	102
Figura 88: Ejemplo del archivo con lista de DOI.	103

Figura 89: Fin de la ejecución de la herramienta de marketing.....103

ÍNDICE DE TABLAS

Tabla 1. Prueba número 1 ejecutada en OSX.....	14
Tabla 2. Fases de la metodología XP y entregables.....	30
Tabla 3. Historias de usuario para desarrollar el Plugin.....	32
Tabla 4: tabla de requisitos para Plugin OJS.....	33
Tabla 5: Pruebas realizadas.....	49
Tabla 6: Historias de usuario para Herramienta de Marketing.....	51
Tabla 7: tabla de requisitos para Herramienta de Marketing.....	52
Tabla 8. Prueba número 1 ejecutada en OSX.....	61
Tabla 9. Prueba número 2 ejecutada en OSX.....	61
Tabla 10. Prueba número 1 ejecutada en Linux distribución Debian7.....	62
Tabla 11. Prueba número 2 ejecutada en Linux distribución Debian7.....	62
Tabla 12. Prueba número 1 ejecutada en Windows7.....	63
Tabla 13. Prueba número 2 ejecutada en Windows7.....	64

RESUMEN

En el presente trabajo se propone el enriquecimiento de artículos científicos alojados en la plataforma Open Journal Systems (OJS) mediante la creación de un plugin, para exponer los denominados article level metrics, el cual extraerá información de dos servicios web específicos; SCOPUS para el número de citas bibliográficas y Altmetric para el número de veces que se comparte un artículo a través de redes sociales y redes de gestión bibliográfica. Además se complementará el enriquecimiento de los artículos ayudando a la difusión de los mismos mediante la creación de una herramienta que complemente a la técnica de e-marketing denominada email marketing.

PALABRAS CLAVE: OJS, SCOPUS, Altmetric, e-marketing, plugin, WebScraping.

ABSTRACT

This thesis proposes the enrichment of scientific articles hosted on Open Journal Systems (OJS) through a plugin development for this platform to expose metrics called “article level metrics”, the plugin will extract information from the SCOPUS and Altmetric web services. SCOPUS to show the number of citations and Altmetric to show how many times an article is shared through social networks and bibliographic management networks. Also, the enrichment of an article will be complemented with the development of a desktop application to help the article diffusion using the e-marketing technique known as email marketing.

KEYWORDS: OJS, SCOPUS, Altmetric, e-marketing, plugin, WebScraping.

INTRODUCCIÓN.

La publicación de datos e información en Internet es un tema que se expande y crece a cada día, conforme las tecnologías, herramientas, técnicas y formatos de visualización se adaptan a las nuevas necesidades de los usuarios en la web, tanto así que (Miniwatts Marketing Group, 2015) registra un crecimiento del 832.5% desde el 2000; de ahí que en la actualidad para que una publicación en un sitio web sea ampliamente difundida no se debe tomar en cuenta únicamente el contenido de la misma, sino también se deben hacer hincapié en otros rasgos como aspecto, accesibilidad, formas de presentación, disponibilidad de los datos, entre otros.

En el caso de publicaciones de artículos científicos y académicos existen sitios y repositorios web que alojan ese tipo de contenidos tales como: PLOS, IEEE, Oxford Journals, cada uno de estos sitios gestiona la información de distintas maneras. De ahí que sea importante para un repositorio de artículos la manera en que administra los contenidos: temas más relevantes, ideas principales, palabras claves y las tecnologías/herramientas que emplea para la organización de la información, así como la manera en que enriquece los datos presentes en un artículo mediante la generación de metadata, la inferencia a partir de información o la aplicación de métricas para obtener información extra.

Debido a ello la presente tesis propone la adaptación de una herramienta conocida como Open Journal Systems (OJS), a necesidades generales de enriquecimiento de contenido mediante técnicas como el WebScraping, todo esto se consolidaría a manera de un plugin que extraerá información del API de SCOPUS así como del servicio de ALTMETRIC para ofrecer al visitante información extra de un artículo en el cual esté interesado, mostrando el impacto que una publicación científica tiene en la web a través de redes sociales como Facebook, Twitter, Google+ y en repositorios que alojan marcadores de publicaciones académicas como Mendeley, CiteULike o Reddit.

Como complemento al enriquecimiento de la información de los artículos, se implementó una aplicación de escritorio desarrollada, la misma que extrae información de las publicaciones/artículos mediante WebScraping, y posteriormente difundir las mismas mediante email marketing¹.

Para llevar a cabo el presente proyecto se ha estructurado la investigación y desarrollo en 6 capítulos.

¹ Técnica de e-marketing que se detallará en el Marco Teórico del presente trabajo.

En el capítulo 1, se define el marco teórico de esta tesis, el software y herramientas que se utilizarán para el manejo de la información y los conceptos básicos que se han contemplado necesarios para el desarrollo del plugin.

Dentro del capítulo 2, se encuentra detallado el Estado del Arte, que es una investigación sobre trabajos afines o similares al presente trabajo; en este apartado se redacta un breve resumen de trabajos que son similares en manejo de información o empleo de tecnologías/herramientas para gestión de contenidos.

En el capítulo 3 se detalla la problemática que da pie a la investigación realizada para este proyecto, así como la solución implementada: metodología utilizada, análisis de información y las propuestas de implementación.

En el capítulo 4 se detalla los dos desarrollos realizados para la presente tesis. El primero sobre el plugin implementado para OJS que además de enriquecer las publicaciones científicas almacenadas dentro del mismo, cambiará la presentación básica de las publicaciones permitiendo resaltar ideas, contenidos o recursos de los mismos que se consideren importantes presentar al público, teniendo como idea base el acceso al artículo desde cualquier dispositivo implementando el diseño responsivo. El segundo desarrollo, detalla la herramienta de marketing científico, la cual realizará una extracción de datos de cada artículo referenciado dentro de una publicación científica en una revista, y posteriormente con esta información construir una base para realizar la difusión masiva de una manera consciente y responsable.

Finalmente en el capítulo 5 se exponen las conclusiones y recomendaciones inferidas gracias al desarrollo del presente trabajo de investigación.

CAPITULO I: MARCO TEÓRICO

1. Conceptos esenciales.

Como parte de la investigación para este proyecto se ha considerado dividir el marco teórico dentro de 3 puntos, acorde a las herramientas y funciones implementadas:

- **Administración de Contenidos.**- debido a que la información se alojará dentro de un repositorio se necesita un marco de referencia para los contenidos, en este apartado se hablará sobre el sistema de manejo de contenidos, los tipos y más específicamente el sistema web de manejo de contenidos.
- **Gestión de la Información.**- en esta sección, que es la más extensa, se describen las herramientas y técnicas empleadas para la publicación y enriquecimiento de la información, así como el software utilizado para la implementación del plugin.
- **Metodologías de Desarrollo.**- en este apartado se da una breve descripción sobre lo que es una metodología de desarrollo, las metodologías ágiles que existen y la metodología usada para el desarrollo de este proyecto.

1.1. Administración de contenidos (CM).

La administración de contenidos se enfoca desde la perspectiva de objetivos y actividades que se le quiere dar a la información que se posee o se quiere recopilar. En la administración de contenidos el objetivo en sí es producir y visualizar publicaciones y documentos digitales. (Boiko, 2001)

El ciclo de vida de la administración de contenidos consta de 3 partes:

- **Colección:** se crea y/o adquiere información.
- **Gestión:** responsable del manejo y almacenamiento de la información, de los usuarios.
- **Publicación:** de manera automática o semiautomática, mediante plantillas, se difunde la información a nivel interno o externo de la organización o empresa que crea los contenidos.

Según (Bluebill Advisors, Inc., 2003) existen 4 macro categorías para la administración de contenidos, de las cuales solo extenderemos el concepto de las dos primeras:

- **CMS:** Content Management System.

- WCMS: Web Content Management System
- ECM: Enterprise Content Management.
- DMS: Document Management System.

1.1.1. CMS: Content Management System.

CMS o Sistema de Gestión de Contenidos, son herramientas empleadas para publicación de contenidos en diferentes formatos y medios a través de la separación de datos y de su formato de presentación (Nakano, 2002). En la publicación (Martín Galán, Rodríguez Mateos, Arellano Pardo, Hernández Pérez, & Nogales Flores, 2005) nos dice que los CMS son programas informáticos que permiten aumentar el volumen, la variedad, la complejidad y el control de los contenidos de un sitio Web; facilitan la interacción entre departamentos de una organización para la publicación de contenidos, que puede ser más dinámica, más actual y estar mejor gestionada.

1.1.2. WCMS: Web Content Management System.

Un WCMS (Sistema Gestión de Contenidos Web, en español) es una herramienta web para la administración de una gran cantidad de recursos, entre ellos: texto, sonido, imágenes, video y demás, mediante herramientas flexibles orientados a la producción de documentos e información digital, especialmente para portales y sitios corporativos. (Robertson, 2003)

Según (Government HKSAR, 2008) mediante un WCMS los usuarios de una empresa/organización pertenecientes a distintos departamentos o ramas dentro de la misma pueden mantener y organizar contenido web en una forma efectiva y sencilla. Además también provee los siguientes beneficios:

- a) Tiempos de respuesta más rápidos, los propietarios del contenido pueden actualizar el material en la web directamente, sin necesidad de asignar esas tareas al personal técnico
- b) Flujos de trabajo más eficientes, mediante un framework WCMS los usuarios de diferentes departamentos pueden modificar y agregar cambios al contenido web mediante un proceso de trabajo predefinido y acordado.

- c) Seguridad mejorada, mediante un framework WCMS, el contenido es únicamente publicado después de su aprobación por gerentes o supervisores designados.
- d) Integración con los servidores de traducción y seguimiento de versiones

Estos sistemas son los más utilizados y erróneamente llamados CMS ya que el ámbito de acción del WCMS es la web y el CMS tiene un rango mucho más amplio e implica todos los contenidos de la organización o empresa.

1.2. Gestión de la información.

“Actualmente, las empresas y organizaciones reconocen la información como un recurso indispensable para ampliar su competitividad, aumentar la calidad y la satisfacción de los clientes, así como para desenvolverse en el mercado global.” (Rojas, 2006)

Es así que toda empresa debe adoptar procesos de gestión de información que le permitan operar, aprender y adaptarse a los cambios de ambiente dentro del ámbito en que se desempeña.

Los procesos principales de la gestión de la información son, según Rojas:

- la identificación de las necesidades de información
- la adquisición de las fuentes informativas
- su organización y almacenamiento
- el desarrollo de productos y servicios
- su distribución y uso

La ejecución de estos procesos conlleva a la creación del conocimiento, al que (Davenport & Prusak, 2001) define: “El conocimiento es una mezcla fluida de experiencia estructurada, valores, información contextual e internacionalización experta que proporciona un marco para la evaluación de nuevas experiencias e información. Se origina y se aplica en la mente de los conocedores. En las organizaciones, con frecuencia no sólo se arraiga en documentos o bases de datos, sino también en las rutinas, procesos, prácticas y normas institucionales”.

Así como la información necesita ser administrada y organizada, el conocimiento al ser un recurso de la misma índole también necesita ser gestionado, puesto que dentro de algunos años una organización sólo podrá mantener ventajas competitivas con respecto a otras mediante el uso adecuado de la información y, más aún, del conocimiento. (Serradell & Juan, 2003)

“Las características del actual ambiente de los negocios exigen que las organizaciones sean cada vez más competitivas, adaptables, flexibles y ágiles. En este entorno, sus estructuras y procesos necesitan renovarse o reestructurarse permanentemente. Las tecnologías de la información, en este sentido, se manifiestan como sistemas ideales para el acceso, búsqueda, tratamiento, utilización y diseminación de información. Funcionan como un mecanismo de soporte para la implementación de nuevos modelos y sistemas de gestión.” (Rojas, 2006)

Es por ello que las organizaciones necesitan modelar e implementar tecnologías y sistemas de información que permitan recopilar, clasificar, administrar y procesar datos para la generación de conocimiento, de esta forma las organizaciones se mantendrán competitivas, ampliarán sus mercados y aumentarán las expectativas de los clientes.

Pero para cumplir con los 3 puntos anteriores una empresa u organización, adicional a un sistema de gestión del conocimiento, del cual se detallará en breve, también necesita de un proceso de difusión, es en este punto que entra el concepto de e-marketing.

1.2.1. Open Access

“Por ‘Open Access’ a la literatura científica se entiende su libre disponibilidad en Internet, permitiendo a cualquier usuario su lectura, descarga, copia, impresión, distribución, búsqueda o enlace a los textos completos de estos artículos, recuperándolos para indexación, pasándolos como datos para software o usándolos para cualquier otro propósito legal, sin barreras financieras, técnicas, legales o de cualquier otra índole.

La única restricción sobre la distribución y reproducción y el único rol para el copyright en este dominio, sería dar a los autores control sobre la integridad de su trabajo y el derecho a que su trabajo sea apropiadamente citado y reconocido.” (Suber, 2012)

El objetivo del Open Acces es maximizar el impacto de los documentos científicos, y esto lo realiza maximizando así mismo el acceso a dichos documentos; esto permite incrementar el acceso a la investigación científica y potenciar el progreso, así como se contribuye al prestigio y reconocimiento de tanto autores como centros de investigación y universidades.

1.2.2. Public Knowledge Project

“El Proyecto de Conocimiento Público o PKP por sus siglas en inglés es un proyecto dedicado a explorar si y como las nuevas tecnologías pueden ser usadas para mejorar el valor público y profesional de la investigación académica.” (Public Knowledge Project, 2010)

Según (Alarcia, 2016) el interés principal del PKP, al igual que el Open Access, es incrementar la accesibilidad hacia los resultados de artículos/publicaciones de investigación académica, y ofrecer también medios de difusión que trasciendan de los ambiente académicos tradicionales, para ello busca elaborar las herramientas informáticas apropiadas y distribuir las gratuitamente.

1.2.3. OJS

Open Journal Systems es una solución de código abierto para la gestión y publicación de revistas académicas en línea. Ofrece un sistema de gran flexibilidad para la gestión y publicación de revistas académicas que puede descargarse sin costo e instalarse en un servidor local, y su funcionamiento queda en manos del equipo editorial de cada institución. El diseño del OJS facilita la reducción del tiempo y la energía que implican las tareas administrativas y de supervisión, propias de la edición de revistas académicas, mejora también la conservación de registros y la eficacia de los procesos editoriales. El objetivo del OJS es mejorar la calidad general y el rigor académico de la publicación de revistas especializadas mediante diversas innovaciones, como mayor transparencia de las políticas que rigen dichas publicaciones y el perfeccionamiento de la indexación. (Public Knowledge Project, 2010)

1.2.3.1. Uso de OJS.

OJS es utilizado en varias revistas en la web. Uno de los retos de la gestión de un proyecto de software de código abierto es que cualquiera puede descargar y utilizar el software sin que se sepa. PKP en su esfuerzo por obtener una aproximación numérica del impacto que su software ha obtenido, ha podido tener información de al menos 24.000 revistas que están utilizando OJS actualmente ya sea en la producción, en una fase inicial, para probar, investigar, experimentar, o para otros fines (por ejemplo, como un sistema de gestión del aprendizaje, como un repositorio de documentos, etc.).

Con el fin de proporcionar una mejor idea de cómo se está utilizando OJS para la publicación de revistas, PKP a través de su página ha publicado una serie de gráficos y mapas centrados específicamente en las revistas que publican por lo menos 10 artículos por año. En la Figura 2 se muestra una perspectiva del impacto de OJS en el año 2015.

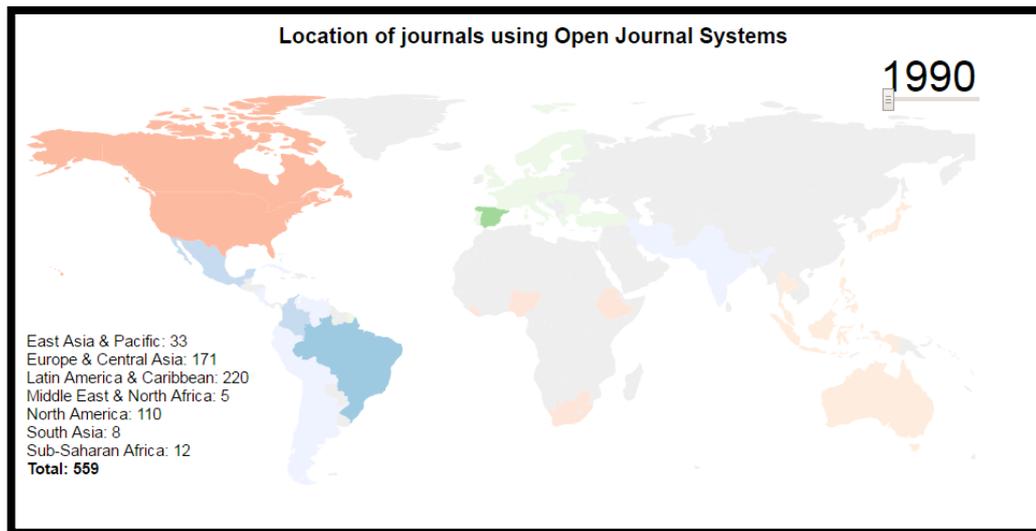


Figura 1: Localización de Revistas usando OJS en el año 1990
Fuente: sitio web oficial de OJS.

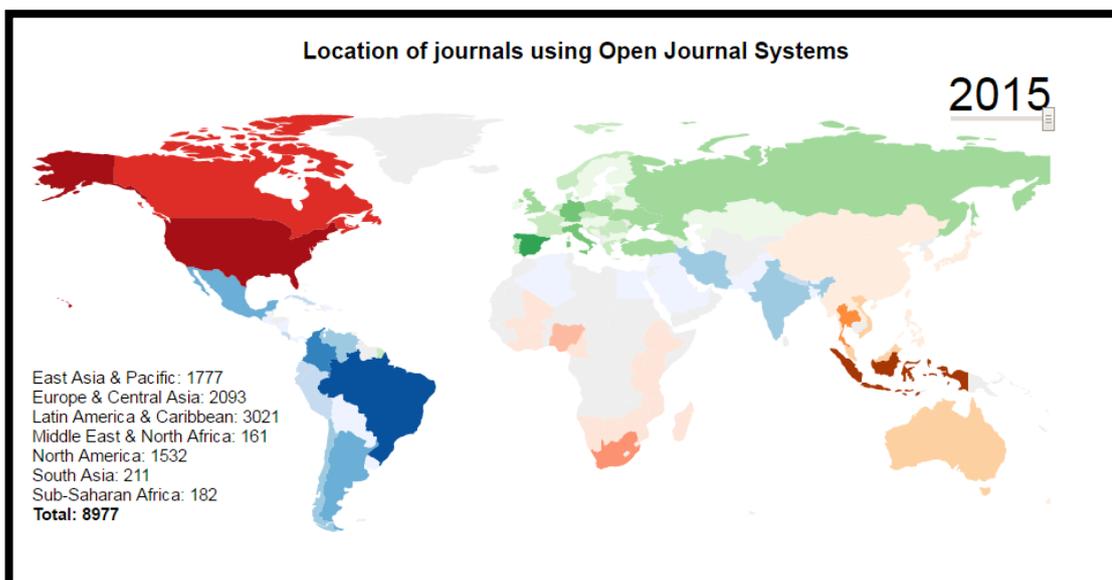


Figura 2: Localización de Revistas usando OJS en el año 2015
Fuente: sitio web oficial de OJS.

En la Figura 2 se puede observar la distribución de las revistas que gestionan sus artículos mediante OJS, el color es un indicativo del grado de expansión de las revistas dentro de una región, siendo los colores más oscuros donde se encuentra la mayor concentración de revistas y los más claros donde existe una cantidad menor de revistas que gestionen sus artículos mediante OJS; si se la compara con la Figura 1 se tiene un indicativo del grado de influencia que está alcanzando OJS dentro del ámbito de publicaciones en línea, conforme han pasado los años el número de revistas que gestiona sus publicaciones mediante OJS ha ido creciendo.

Entre las gráficas que OJS nos ofrece, también podemos observar diagramas de barras del crecimiento y la acogida que ha tenido esta herramienta desde el año 1990 mediante 3 parámetros diferentes:

- Número de revistas que la usan (Figura 3).

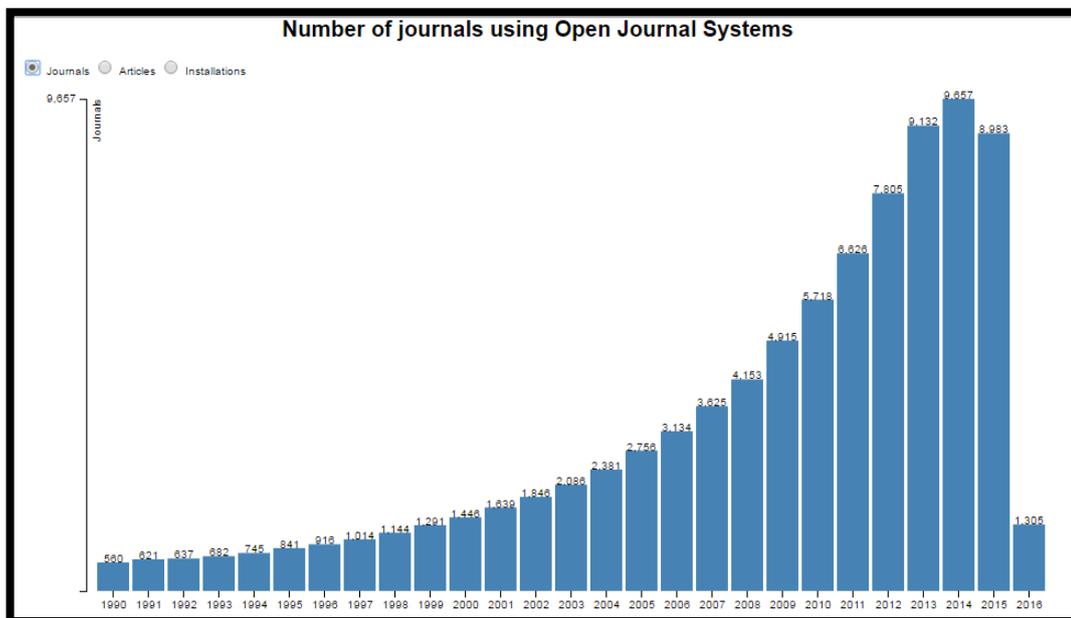


Figura 3: Gráfico que representa el crecimiento del uso de OJS desde 1990 hasta el año actual
Fuente: sitio web oficial de OJS. <https://pkp.sfu.ca/ojs/ojs-usage/ojs-map/>

- Número de artículos publicados año por año.

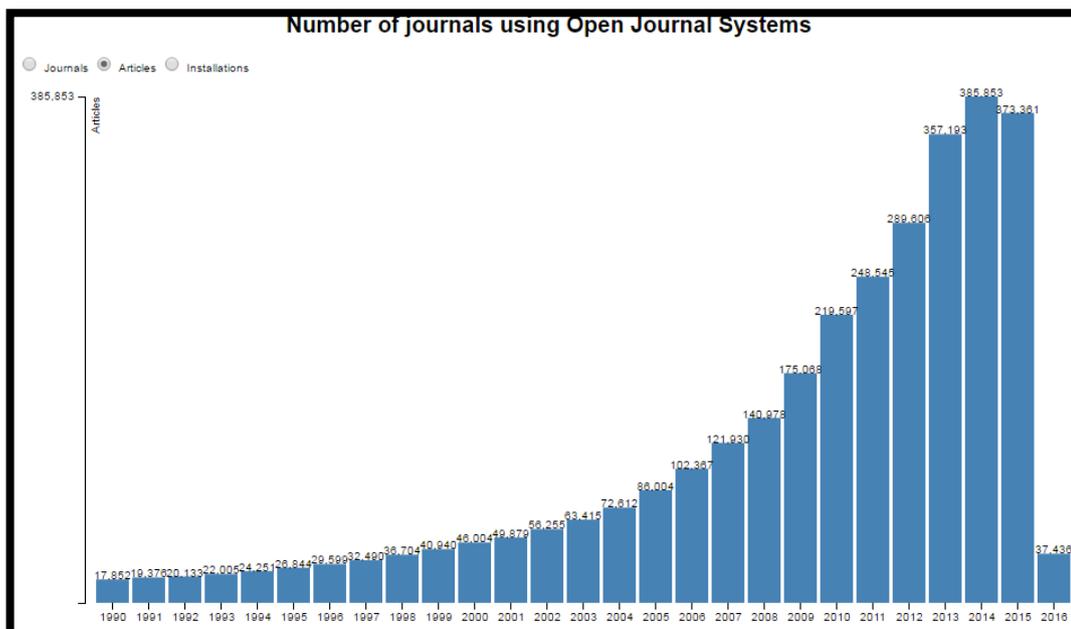


Figura 4: Gráfico que representa el crecimiento del número de artículos publicados mediante OJS desde 1990.
Fuente: sitio web oficial de OJS. <https://pkp.sfu.ca/ojs/ojs-usage/ojs-map/>

- Número de servidores que alojan la herramienta.

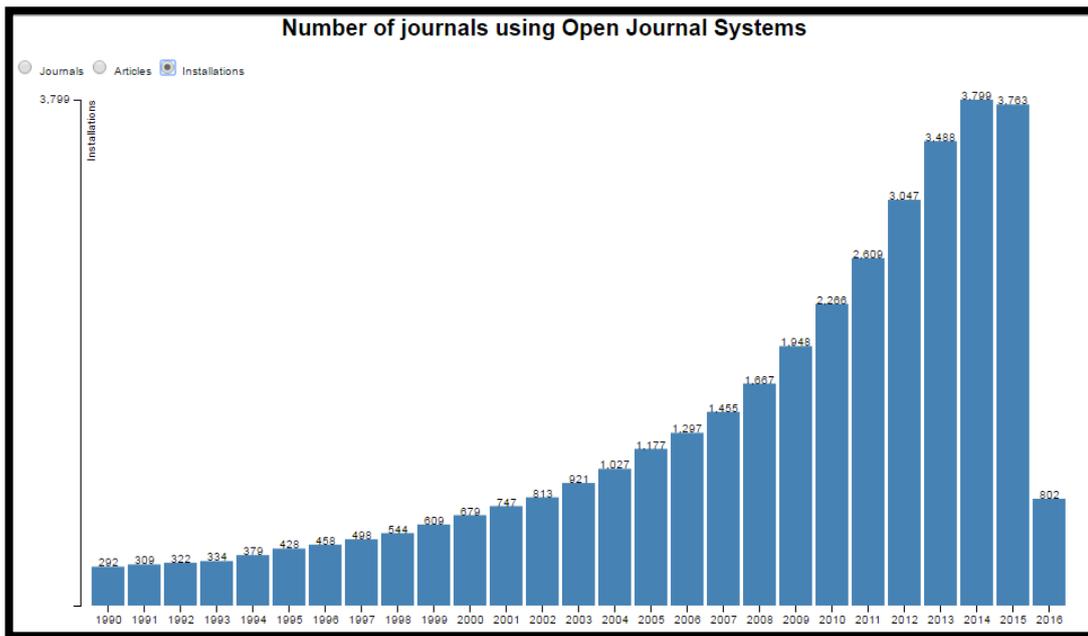


Figura 5: Gráfico que representa el crecimiento del número de instalaciones desde 1990 que hacen uso de OJS para sus publicaciones.

Fuente: sitio web oficial de OJS. <https://pkp.sfu.ca/ojs/ojs-usage/ojs-map/>

1.2.3.2. Casos de éxito implementando OJS en Sudamérica.

En Colombia, esta herramienta es utilizada por la UIS, Universidad del Rosario, Universidad Nacional, la Pontificia Universidad Javeriana, la Universidad Icesi, la Universidad de Antioquia y la Universidad de Cundinamarca, entre otras.

En Argentina, el CAICYT de CONICET está implementando OJS en el Portal de Publicaciones Científicas y Técnicas para revistas científicas argentinas, y la Universidad Nacional de La Plata posee desde 2008 el Portal de Revistas.

En Ecuador, la Escuela Politécnica Nacional a través del Departamento de informática y Ciencias de la Computación utiliza Open Journal Systems para su revista electrónica ReDiFIS. <http://redifis.epn.edu.ec>, y la Universidad Católica de Santiago de Guayaquil, usa OJS en su publicación de Revista Medicina.

1.2.4. E –Marketing

Según (Juslén, 2011) para que una empresa se mantenga competitiva en el mercado laboral actual y obtenga mejores resultados no basta con utilizar los medios de marketing tradicional, es necesario emplear aplicaciones y estrategias de marketing originadas en el Internet, a partir de todos aquellos canales potenciales que pueden ser usados para el marketing, y enfocadas en el marketing online.

Esto se conoce como Marketing Electrónico, E-Marketing por sus siglas en inglés, y (Pride & Ferrell, 2012) lo define como: “el proceso estratégico de distribución, promoción, fijación de precios y descubrimiento de los deseos de los clientes usando medios digitales y marketing digital. El E-Marketing va más allá del Internet y también incluye teléfonos móviles, banners publicitarios, marketing digital al aire libre², y las redes sociales”.

Entre las técnicas para el E-Marketing se tiene (Oulu, 2012):

- a) Banners, es una pantalla gráfica la cual es enlazada a los anunciantes.
- b) E-mail marketing o publicidad e-mail, publicidad específica enviada a una lista de mails de un grupo de usuarios seleccionado.
- c) Publicidad mediante motores de búsqueda, las compañías de publicidad pagan a los motores de búsqueda para que muestre publicidad en las búsquedas de los usuarios, según palabras claves que definan las empresas.
- d) Posicionamiento en buscadores, técnica mediante la cual se incrementa el rango de un sitio en los motores de búsqueda, aplicando algoritmos y frases de búsqueda
- e) Análisis en la web, se rastrean la forma en que los usuarios recorren un sitio web, para así optimizar la navegación.

Se explicará más detalladamente el método que se empleará para este proyecto:

1.2.4.1. E-mail Marketing.

“Como una forma de marketing digital el envío de información de una compañía o un producto a gente que se encuentre en una lista de correo, se ha convertido en una forma popular de publicidad en el Internet y además es una importante fuente de tráfico. Tradicionalmente esto es hecho al pedir a un consumidor que llene una indicando sus

² Vallas publicitarias digitales o impresas.

intereses cuando se registran para un servicio, entonces el comercializador entrelazan los mensajes de publicidad con los intereses de los consumidores.” (Turban, King, Lee, Liang, & Turban, 2010)

Según (Turban, King, Lee, Liang, & Turban, 2010) la ventaja de este medio es que los mails:

- son medios interactivos que combinan publicidad y servicio al cliente
- tienen un costo relativamente bajo frente a otros medios de marketing
- permiten enfocarse en un audiencia específica

Así mismo gracias a las bases de datos que la mayoría de organizaciones/compañías manejan se puede realizar marketing digital de forma masiva y hasta específica según el tipo de audiencia que se identifique.

1.2.4.2. Search Engine Optimization

“Cuando Google apareció en 1998, las personas descubrieron que proveía mejores resultados que otros motores de búsqueda; la razón de este fenómeno se debía en gran parte gracias al algoritmo de popularidad de los enlaces que Google poseía. Progresivamente las personas dedicadas al marketing aprendieron tanto como pudieron acerca de estos algoritmos y optimizaron sus sitios web adecuadamente; a partir de ello, metodologías para la optimización de sitios en motores de búsqueda comenzaron a aparecer y aun son relevantes.” (Sweeney, MacLellan, & Dorey , 2006)

La optimización de motores de búsqueda o SEO por sus siglas en inglés es un método empleado para incrementar el rango de un sitio en los motores de búsqueda. Para ello el método necesita conocer el algoritmo de ranking del buscador y las frases más adecuadas de búsqueda, y posteriormente adapta la publicidad en consecuencia. (Turban, Lee, King, Liang , & Turban, 2010)

1.2.4.3. Article Level Metrics

Según (Fenner, 2013): “Article Level Metrics provee un amplio rango de métricas acerca de la recepción de un artículo individual de revista por la comunidad científica después de su publicación. Incluyen citas, estadísticas de uso, discusiones en comentarios en línea y redes sociales, marcadores sociales y recomendaciones.”

De ahí se puede afirmar que mediante la aplicación de ALMs se puede obtener un conjunto de medidores sobre el alcance que tiene un artículo en una revista, y también la interacción que el mismo ha generado dentro de un ámbito específico.

1.2.5. Web Scraping

El web scraping es una técnica que en los últimos años ha sido cada vez más utilizada con el fin de obtener de una manera mucho más rápida datos estructurados de una fuente no estructurada como lo es la web. Lo cierto es que el web scraping, aunque es una técnica que ahorra mucho tiempo de procesamiento, es tediosa y frágil; es decir que la efectividad de una herramienta, creada mediante esta técnica, está muy estrictamente sujeta al posible cambio que el administrador de la página web realice (de la cual se quiere obtener los datos). (Penman, Baldwin, & Martinez, 2009)

La técnica del web scraping consiste en la simulación de la navegación que una persona pudiera realizar en cierto sitio web mediante un programa automático que procesará el contenido no estructurado, por lo general HTML, para almacenar ciertos datos que consideremos nos puedan servir para algún propósito. (Hernández, 2014)

1.2.6. DOI.

Digital Object Identifier o identificador digital de objeto en español, es un sistema de identificación único para un objeto que puede ser cualquier entidad o cosa física, digital o abstracta que apareció en el año de 1997 por la iniciativa de tres asociaciones dedicadas a la publicación. Luego en el mismo año se creó la International DOI Foundation (IDF) para darle continuidad a dicho proyecto. (DOI®, 2014)

En 2000 se lanzó el primer sistema de citado de artículos electrónicos por el CrossRef Registration Agency; y en 2010 el sistema DOI fue aprobado como un estándar ISO, el estándar asignado es ISO 26324. (DOI®, 2014)

1.2.6.1. Sintaxis de nombramiento DOI.



Figura 6: Estructura del DOI.

Fuente: Cite my data policy statement. <http://www.ands.org.au/>

Consiste en una secuencia alfanumérica única que contiene dos partes separadas del carácter barra inclinada “/”: (DOI®, 2014)

- Publisher ID: indica el número que le asigna la Agencia DOI al editor u entidad que lo requiera. (Prefijo)
- Item ID: es un identificador que le asigna el editor concreto y que puede ser una secuencia alfanumérica de caracteres. (Sufijo)

Esta asignación de prefijo (Publisher ID) a las entidades que lo requieran, evita una necesidad de nombramiento centralizado, solo se debe tener en cuenta que para cada DOI debe existir una combinación de prefijo y sufijo (Item ID) único.

1.2.6.2. Ventajas del sistema DOI.

El sistema DOI ofrece un conjunto único de funcionalidades entre las cuales destacan: (DOI®, 2014)

- Persistencia, si se mueve el material, reorganizado o remarcado;
- Interoperabilidad con otros datos procedentes de otras fuentes;
- Gestión de datos única para múltiples formatos de salida (independencia de plataforma);
- Actualización dinámica de metadatos, aplicaciones y servicios.

Los beneficios de la implementación del sistema DOI incluyen facilitar la gestión de contenidos internos y el desarrollo de productos más escalables.

1.3. Metodologías de Desarrollo.

“Una metodología es un conjunto de procedimientos que se estructuran y realizan con la función de conseguir unos objetivos.” (García, 2015)

Dado que el desarrollo de software es una tarea que en ocasiones resulta sino bien difícil si bastante compleja, es por ello que existen numerosas propuestas metodológicas enfocadas en distintas dimensiones del proceso de desarrollo; existen aquellas metodologías tradicionales en que se controlan las actividades involucradas, los artefactos resultantes y las herramientas empleadas, dichas metodologías han probado ser efectivas y necesarias para el desarrollo de varios proyectos, pero así mismo han presentado problemas en otros muchos; para evitar la complejidad añadida al resolver estos problemas adicionando más controles dentro de estas metodologías se cambió el enfoque de desarrollo centrándose más en el factor humano o en el producto software, esta es la filosofía que adoptan las metodologías ágiles. (Canós, Letelier, & Penadés, 2003)

Dado el ámbito de este proyecto se empleará una metodología ágil para el desarrollo del mismo, a continuación describimos las principales características de una metodología ágil y las diferencias que existen entre estas y las metodologías tradicionales.

1.3.1. Metodologías Ágiles.

“El gran cambio que supone la aproximación ágil radica en que en lugar del equipo de trabajo basarse en procesos y herramientas, como se hacía en la metodología tradicional, ahora se le da más importancia a la forma de trabajar de los propios miembros del equipo y a la forma que estos tienen de realizar las tareas.” (García, 2015)

Una de las diferencias más notables de las metodologías ágiles contra las tradicionales es que en lugar de centrarse en procesos y generación de documentación, se le da un mayor valor al individuo, a la colaboración con el cliente y al desarrollo incremental del software con iteraciones muy cortas. (Canós, Letelier, & Penadés, 2003)

A continuación se presenta una tabla en la que se exponen las diferencias más significativas entre una metodología ágil y una tradicional (“no ágil”).

Tabla 1. Prueba número 1 ejecutada en OSX

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

Fuente: tomado de (Canós, Letelier, & Penadés, 2003)

Aunque las metodologías ágiles se atienen a ciertos principios, cada una de ellas atiende a distintas pautas y posee sus propias características. A continuación se detallan 4 metodologías ágiles (Canós, Letelier, & Penadés, 2003):

- **SCRUM.-** Desarrollada por Ken Schwaber, Jeff Sutherland y Mike Beedle. Define un marco para la gestión de proyectos, que se ha utilizado con éxito durante los últimos 10 años. Está especialmente indicada para proyectos con un rápido cambio de requisitos. Sus principales características se pueden resumir en dos. El desarrollo de software se realiza mediante iteraciones, denominadas sprints, con una duración de 30 días. El resultado de cada sprint es un incremento ejecutable que se muestra al cliente. La segunda característica importante son las reuniones a lo largo proyecto, entre ellas destaca la reunión diaria de 15 minutos del equipo de desarrollo para coordinación e integración.
- **Crystal Methodologies.-** Se trata de un conjunto de metodologías para el desarrollo de software caracterizadas por estar centradas en las personas que componen el equipo y la reducción al máximo del número de artefactos producidos. Han sido desarrolladas por Alistair Cockburn. El desarrollo de software se considera un juego cooperativo de invención y comunicación, limitado por los recursos a utilizar. El equipo de desarrollo es un factor clave, por lo que se deben invertir esfuerzos en mejorar sus habilidades y destrezas, así como tener políticas de trabajo en equipo

definidas. Estas políticas dependerán del tamaño del equipo, estableciéndose una clasificación por colores, por ejemplo Crystal Clear (3 a 8 miembros) y Crystal Orange (25 a 50 miembros).

- **Dynamic Systems Development Method (DSDM).**- Define el marco para desarrollar un proceso de producción de software. Nace en 1994 con el objetivo de crear una metodología RAD unificada. Sus principales características son: es un proceso iterativo e incremental y el equipo de desarrollo y el usuario trabajan juntos. Propone cinco fases: estudio viabilidad, estudio del negocio, modelado funcional, diseño y construcción, y finalmente implementación. Las tres últimas son iterativas, además de existir realimentación a todas las fases.
- **Adaptive Software Development⁸ (ASD)** [9]. Su impulsor es Jim Highsmith. Sus principales características son: iterativo, orientado a los componentes software más que a las tareas y tolerante a los cambios. El ciclo de vida que propone tiene tres fases esenciales: especulación, colaboración y aprendizaje. En la primera de ellas se inicia el proyecto y se planifican las características del software; en la segunda desarrollan las características y finalmente en la tercera se revisa su calidad, y se entrega al cliente. La revisión de los componentes sirve para aprender de los errores y volver a iniciar el ciclo de desarrollo.

Finalmente se encuentra la más representativa de las metodologías ágiles: la metodología XP, que será también usada para el desarrollo de este proyecto y la cual se detallará en el siguiente apartado.

1.3.2. Metodología XP.

“XP es una metodología ágil centrada en potenciar las relaciones interpersonales como clave para el éxito en desarrollo de software, promoviendo el trabajo en equipo, preocupándose por el aprendizaje de los desarrolladores, y propiciando un buen clima de trabajo. XP se basa en realimentación continua entre el cliente y el equipo de desarrollo, comunicación fluida entre todos los participantes, simplicidad en las soluciones implementadas y coraje para enfrentar los cambios. XP se define como especialmente adecuada para proyectos con requisitos imprecisos y muy cambiantes, y donde existe un alto riesgo técnico.” (Canós, Letelier, & Penadés, 2003)

CAPITULO II: ESTADO DEL ARTE

2. Trabajos relacionados sobre enriquecimiento de información.

A continuación se detallarán trabajos que se encuentran relacionados ya sea mediante la implementación de métricas o el enriquecimiento de información de los artículos científicos mediante distintas herramientas o tecnologías.

2.1. Altmetric: enriching scholarly content with article-level discussion and metrics.

En este proyecto se ha realizado un análisis sobre Altmetric como métrica para calcular el impacto que tiene un artículo en línea, se presenta dentro del mismo estudio la siguiente tabla de datos:

Service	14–21 June	6–13 September	% change
Twitter	37,858	43,542	+13
Facebook	1,216	3,181	+62
Reddit	156	154	-1
Pinterest	40	133	+69
Blogs	471	430	-10

Figura 7: Volumen de actividad de servicios en línea que involucran un enlace a un artículo académico, rastreado por Altmetric
Fuente: tomado de (Adie & Roe, 2013)

En dicha tabla se puede apreciar el crecimiento que han tenido un determinado conjunto de artículos en un lapso de aproximadamente 8 semanas, todo ello calculado mediante la herramienta Altmetrics, dicho porcentaje de crecimiento se puede observar en la última columna.

Dentro del proyecto se menciona el crecimiento que puede tener un artículo en las redes sociales, sea mediante referencias, retuits, discusiones, veces compartidas, likes, entre otras interacciones con el mismo; se menciona para este caso el crecimiento que tuvo el artículo principal del consorcio ENCODE publicado en septiembre del 2012

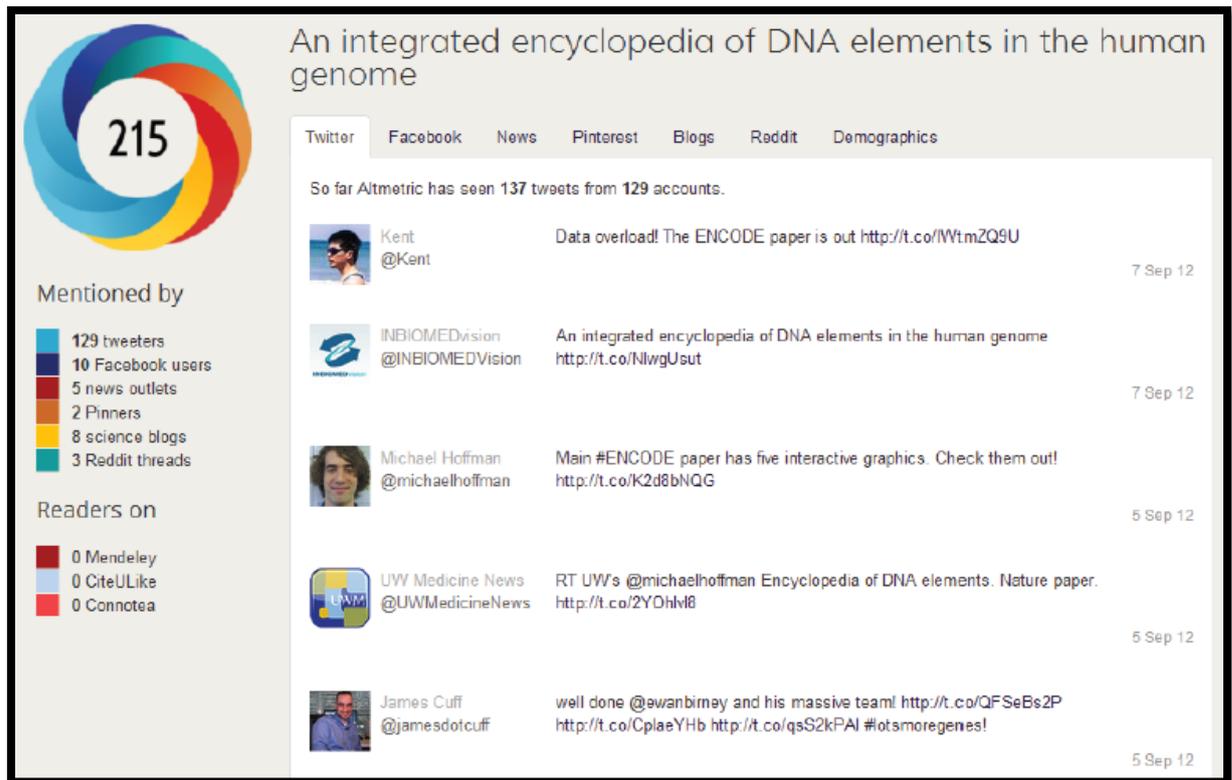


Figura 8: detalles de Altmetric para el paper ENCODE.
Fuente: tomado de (Adie & Roe, 2013)

En la Figura 8 se puede observar el uso de la herramienta Altmetrics para recoger la difusión dentro de las redes sociales en el caso específico del paper “An integrated encyclopedia of DNA elements in the human genome”; se puede apreciar el puntaje Altmetric del artículo, generado por la herramienta Altmetrics, así como las menciones y los lectores que lo han revisado.

También se menciona en el estudio el crecimiento que el número de artículos únicos que se discuten, comparten y/o visualizan tiene un crecimiento mensual del 5% al 10% (Adie & Roe, 2013)

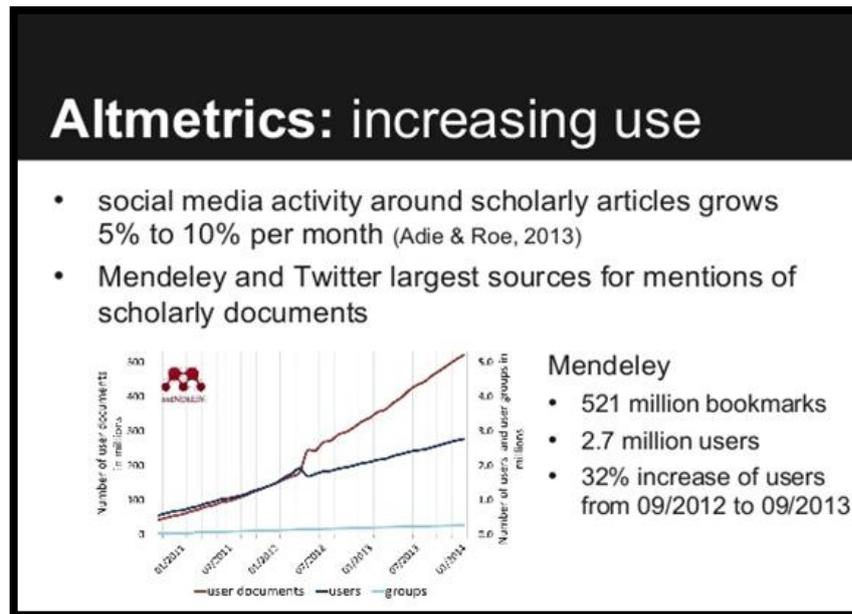


Figura 9: Gráfica del incremento de uso de Altmetric.
Fuente: Diapositiva 13 de la siguiente presentación <http://goo.gl/dQzHiJ>.

2.2. A graphical article-level metric for intuitive comparison of large-scale literatures

En este artículo publicado por (Ling, Liu, Huang, Shah , & Li, 2015) se propone un métrica a nivel de artículo que sea gráfico (gALM por sus siglas en inglés), el cual capture el impacto y la popularidad de los artículos académicos desde los aspectos científicos y sociales. Al capturar estas 2 dimensiones se grafica un mapa circular, el cual se encuentra dividido en sectores de documentos perteneciendo a un año de publicación y cada bloque representa citas de revistas de artículos, por tamaño de bloque, y lectores en Mendeley por el color del bloque. De esta manera gALM provee una comparación más intuitiva de literaturas de larga escala.

Así mismo se ha diseñado el Science Navigation Map (SNM), el cual es un servidor web online que permite no solo visualizar las gALM sino que las provee de rasgos interactivos, de esta manera los usuarios pueden intuitivamente hacer comparaciones de artículos así como explorar y filtrar información importante y relevante para estas publicaciones, todo esto mediante un mapa de visualización interactivo de métricas a nivel de artículo, dichas métricas miden tanto el impacto científico como la popularidad social en Mendeley,. A continuación se muestra una visualización de todos los artículos publicados en PLoS Biology durante el 2003 y el 2014 por SNM.

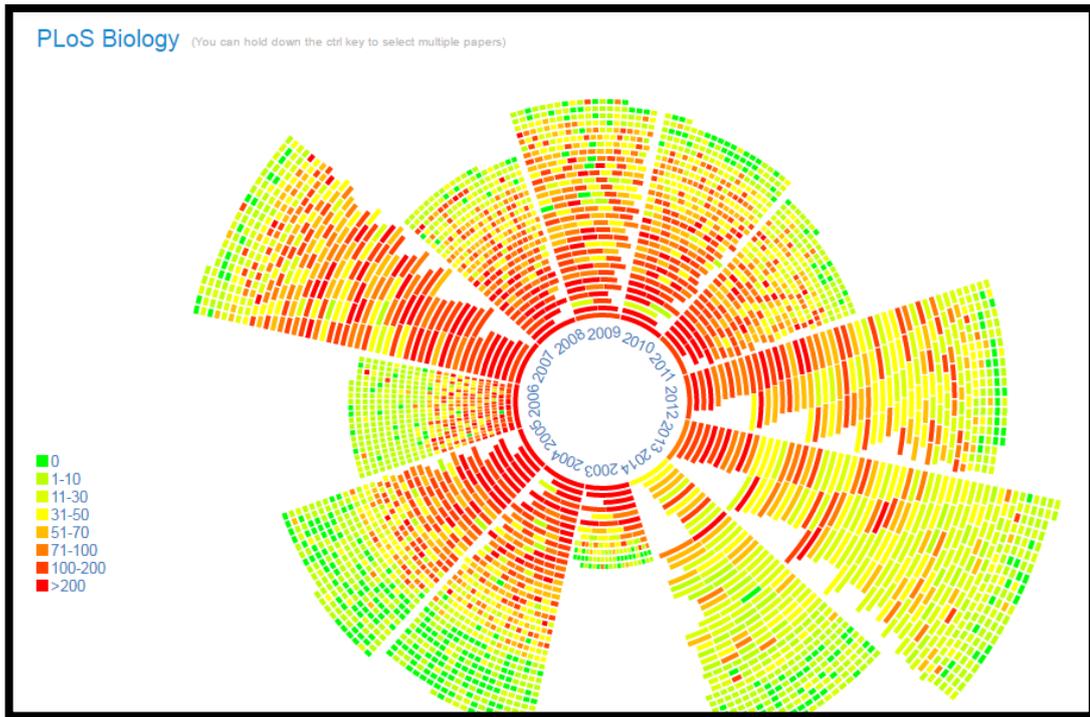


Figura 10: Papers de PLoS Biology publicados durante 2003 y 2014 por SNM .
Fuente: tomado de <http://mirror2.linkscholar.org/plosbiology/>

Mediante este mapa se puede encontrar estadísticas básicas de artículos tales como los más citados y los más populares en Mendeley durante un período de tiempo.

2.3. Métodos y técnicas para el cálculo y visualización de métricas de historiales en artículos de wikis.

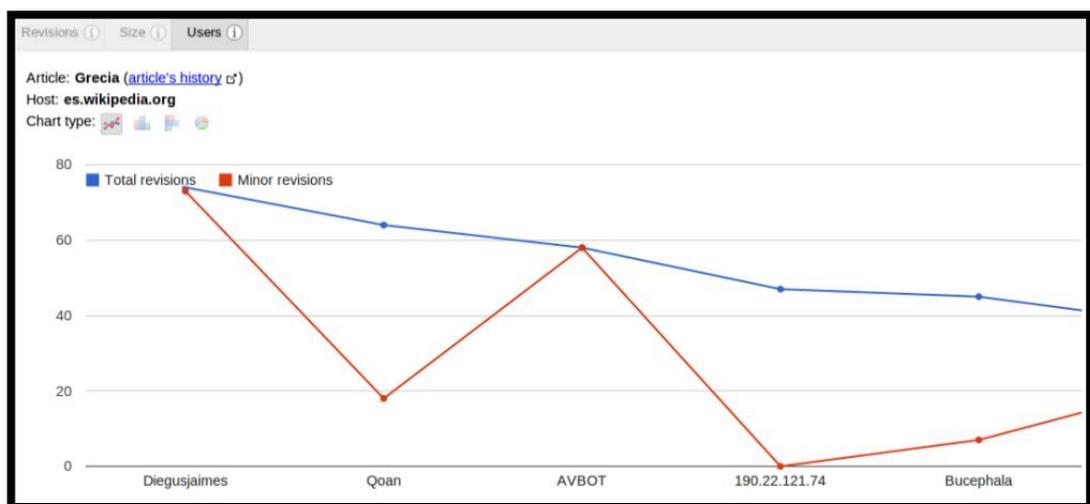


Figura 11: Métrica de revisiones por usuario.
Fuente: Tomado de (D'APUZZO & WORWA, 2012)

Los wikis mantienen un registro de los cambios o ediciones que ha sufrido un artículo a través del tiempo, entre los cuales se detallan diversos atributos como: los usuarios que le han editado, las fechas y el tamaño de las ediciones, entre otros. A pesar que dicha información es útil, muchos de los wikis no aprovechan las propiedades presentes en la misma. Por ejemplo, se puede tener una lista con los cambios realizados a un artículo, pero no existe manera de saber, a simple vista, qué usuario ha contribuido más en un artículo dado o incluso en qué fecha un artículo ha sufrido más ediciones. De igual manera, no se cuenta con visualizaciones gráficas que pueden resultar de gran utilidad a la hora de consultar y contrastar cambios variantes en el tiempo. Con la creación de una herramienta automatizada, se recupera la información presente en el historial de los artículos de un wiki de manera periódica, para el posterior cálculo de métricas que luego serán visualizadas de manera gráfica. (D'APUZZO & WORWA, 2012)

CAPITULO III: PROBLEMÁTICA Y SOLUCIÓN PROPUESTA

3. Problemática y Objetivos.

3.1. Problemática actual.

En la actualidad la gestión de artículos científicos llevada a cabo mediante OJS muestra información como el título, el DOI, el resumen, y un enlace al texto completo del artículo, debido a esto los únicos datos relevantes que se puede obtener era la información mostrada, la inferencia de conocimiento queda atada al plano personal del usuario, y la tarea de intentar medir el impacto que un artículo en específico había tenido dentro de la comunidad académica/científica resulta en un proceso bastante extensa y difícil.

De ahí que el problema en la gestión de publicaciones en OJS es que no existe una forma de medir el impacto de un artículo en los usuarios o consumidores, es aquí donde entra la influencia de las redes sociales en la vida de los usuarios y los repositorios de información en el ciclo de vida de un artículo.

Según (Porrúa García, 2010): “Las redes sociales a través de Internet únicamente han supuesto una extensión de lo que todos venimos haciendo desde que nacemos: relacionarnos con nuestra familia, posteriormente con nuestros amigos y a continuación con nuestro entorno profesional, conformando así a nuestro alrededor una red de contactos con la que de modo más o menos intermitente vamos interactuando en el transcurso de nuestra existencia.”

De esto se puede afirmar que en la actualidad para medir el impacto de una noticia, un post, un vídeo, un documento o más aún, en este contexto, una publicación se debe examinar y analizar la interacción que el mismo ha generado no sólo dentro de las redes sociales, sino también en repositorios de información y la retroalimentación que ha habido a partir del mismo.

Para mejorar la interacción de los usuarios con las publicaciones es necesario incrementar el tráfico (visitas) hacia el repositorio OJS en que se encuentren publicadas, para ello es necesario optimizar el proceso de difusión de los artículos publicados en OJS, pero la carencia de una herramienta especializada para obtener información de personas que estén potencialmente interesadas en los temas publicados hacen difícil la difusión hacia un público interesado en temas puntuales, artículos específicos o contextos especializados dentro de OJS.

Una técnica para este proceso de difusión consiste en la creación de un método de suscripción para que las personas interesadas obtengan correos electrónicos cada vez que un nuevo artículo es publicado (Pride & Ferrell, 2012), pero esto no resuelve el problema de la generación inicial de tráfico hacia el nuevo artículo publicado. Algunas revistas recurren al envío masivo de correos electrónicos sin una suscripción previa como lo es el caso de universidades que envían a sus estudiantes información de nuevas publicaciones de sus revistas, el inconveniente con esta técnica es el problema de crear SPAM masivo y no focalizado.

3.1.1. Propuesta para OJS.

OJS actualmente presenta la información sobre los artículos mediante texto plano, resultando esta visualización monótona y simple, sin ningún dato adicional que aporte valor agregado al lector del artículo, tal y como se muestra en la Figura 12:



Figura 12: visualización de un artículo con la actual visualización en OJS.
Fuente: Autor

Otro problema que se encuentra en dichas revistas es que no existe una forma de enriquecimiento de los datos a partir de la interacción o la influencia de los mismos en las investigaciones de los usuarios, no existe una métrica mediante la recopilación de interacciones en las redes sociales, es decir que no se les puede dar valor agregado a las publicaciones y la difusión o penetración que han tenido en las redes sociales no se vería reflejada en el origen de la misma.

Debido a la nueva revisión del lenguaje básico de la World Wide Web llamado HTML5³, se ha obligado prácticamente a muchos sistemas enfocados a la web, en especial manejadores de contenido, a actualizar su código para no caer en la obsolescencia y sigan captando público en general y aunque los navegadores más populares (Chrome, Firefox, Safari, etc.) aún mantienen la capacidad de interpretar páginas web con versiones anteriores de HTML en sus diseños, claramente la tendencia está fuertemente orientada al cambio hacia el nuevo estándar del lenguaje debido a las nuevas etiquetas que harán un cambio significativo a nivel de visualización de tipos de archivos que no estaban soportados sin necesidad de cargar aplicaciones adicionales como audio y/o video; otra de las ventajas de cambiarse al estándar HTML5 es que ayudará al posicionamiento de las páginas ya que al hacer una búsqueda en Google se priorizarán las que estén escritas en este nuevo estándar frente a las anteriores.

Se tiene como antecedente que Open Journal Systems es un sistema creado en el lenguaje PHP y orientado a objetos, que además utiliza plantillas HTML manejadas con Smarty para la interfaz de usuario y pese a que sus plantillas están declaradas con el estándar XHTML⁴ en su normativa XHTML-1.0-Transitional se tratará de modificar las plantillas para incluir HTML5 y así poder obtener, mediante la incorporación de BootStrap, un efecto denominado Responsive Design (o diseño responsivo en español) para que se pueda visualizar los artículos a través de las pantallas de distintos dispositivos móviles sin recurrir a modificaciones puntuales en el código.

3.1.2. Propuesta para marketing.

Puesto que recibir correos que no deseamos puede ser muy incómodo y muchas de las veces resulta en bloqueos por parte de quien los recibe, se cree que una mejor estrategia para la difusión de nuevos números de la revista en línea sería enfocarlos a gente que esté relacionada con los temas publicados.

Cada artículo publicado de una revista científica, posee en la parte final del mismo un conjunto de citas llamadas “referencias” que sustentan la investigación realizada o dan un entendimiento más claro extendiendo temas que se han tratado dentro del artículo en cuestión. La mayoría de artículos de una investigación científica va a estar apoyada por al menos una referencia a otra investigación científica en la misma línea ya sea publicada en la

³HyperText Markup Language, versión 5

⁴eXtensible Hyper Text Markup Language

misma revista o en alguna otra; estas revistas científicas utilizan los DOI, ya que es un identificador único, para ubicar puntualmente a un artículo. Teniendo esto como premisa, se ha visto la oportunidad de aprovechar la información que existe en estas referencias para que mediante el aprovechamiento del API de SCOPUS con su recurso de búsqueda a través de DOI, obtener información adicional de los artículos referenciados que en este caso sería la obtención del email del autor principal y así construir una base de datos de personas interesadas en las publicaciones de la revista publicada.

3.2. Objetivos.

General:

- Desarrollar un plugin para el enriquecimiento de la información de un artículo publicado a través de la plataforma OJS (Open Journal Systems).
- Crear una aplicación de escritorio que mediante técnicas de web scraping obtenga información relevante para poder realizar marketing científico.

Específicos:

- Implementar una nueva plantilla que permita reorganizar y reestructurar la información presentada a un usuario sobre los artículos alojados en un repositorio OJS.
- Enriquecer la información de artículos científicos mediante la implementación de métricas recopiladas a través de Altmetric y Scopus.
- Documentar los procesos de análisis y desarrollo para establecer un marco de trabajo para futuros desarrollos de plugins para OJS.
- Implementar una aplicación de escritorio que permita extraer el correo, el título y otros datos de un artículo para mediante e-mail marketing se pueda realizar la difusión focalizada de las publicaciones a un grupo específico de usuarios.

CAPITULO IV: DESARROLLO E IMPLEMENTACIÓN

4. Desarrollo e Implementación.

4.1. Propósito.

El enriquecimiento de un artículo mediante un plugin de OJS es uno de los objetivos principales de la presente tesis, a través de este enriquecimiento el lector de los artículos publicados podrá calcular el impacto que ha tenido la publicación en una red social o en un repositorio específico esto se lo realizará mediante la implementación de métricas.

Como parte del proceso de enriquecimiento también se implementará una aplicación de escritorio que permita extraer título y correo de un artículo para posteriormente mediante email marketing realizar la difusión del mismo, esto con el fin de incrementar el tráfico hacia los repositorios OJS e incrementar la interacción de los usuarios con las publicaciones, lo que a su vez generarán métricas que permitan seguir alimentando el proceso de enriquecimiento.

4.2. Metodología.

Para llevar a cabo el desarrollo del plugin, se ha propuesto el uso de la metodología ágil llamada programación extrema o XP por sus siglas en inglés. Se considera que esta metodología es la apropiada para trabajar en el desarrollo del mismo, puesto que los requerimientos no están ligados a un cliente en particular, se necesitará realizar mejoras puntuales y de forma rápida e inmediata, y los requerimientos estarán sujetos a cambios durante el desarrollo, además la metodología XP se centra en el desarrollo y la retroalimentación obtenida a partir de los requerimientos implementados y sobre las pruebas unitarias del mismo.

Aplicando esta metodología se obtendrán los siguientes entregables:

Tabla 2. Fases de la metodología XP y entregables.

Fase	Tareas	Documentos
Planificación	a) Identificar en el mundo real los objetos. b) Reuniones de seguimiento diarias. c) Correcciones, las partes involucradas deben estar conscientes de los cambios realizados.	<ul style="list-style-type: none">Historias de usuario
Diseño	a) Verificar si el diseño satisface todos los requisitos identificados mediante iteraciones con el usuario.	<ul style="list-style-type: none">Iteraciones
Desarrollo	a) Implementación de requisitos y cambios en código fuente. b) Integración del código. c) Realizar reciclaje de código que implicará mantener el código limpio y fácil de comprender, modificar y ampliar.	<ul style="list-style-type: none">Código generado
Pruebas	a) Comprobar el correcto desempeño de las funcionalidades que se han implementado b) Realizar pruebas.	<ul style="list-style-type: none">Pruebas

Fuente: Autor.

4.3. Desarrollo del plugin.

Para el desarrollo e implementación del plugin en OJS se modela con la siguiente secuencia de pasos:

- a) Planificación.- se toman las historias de usuario para definir los requisitos y el modelado del sistema.
- b) Diseño.- en esta fase se identifican los requisitos y se mantienen reuniones con el usuario para mostrar las implementaciones realizadas, esto con el fin de obtener retroalimentación por parte del usuario para implementar nuevos requisitos o modificar los ya existentes.
- c) Desarrollo.- en esta fase se implementan los requisitos en código fuente.
- d) Pruebas.- se establecen pruebas de funcionalidad para verificar el correcto funcionamiento de los requisitos implementados y si satisfacen los requerimientos del usuario.

4.3.1. Planificación.

Se toman las historias de usuario para obtener un bosquejo de las distintas funcionalidades y cambios requeridos según la interacción con el usuario y la retroalimentación obtenida.

4.3.1.1. Historias de usuario

Descripciones claras por parte del cliente, sin entrar en detalles de la manera de codificar, estas historias de usuario nos permiten obtener una idea de los requerimientos que el cliente desea satisfacer mediante la implementación del plugin.

A continuación en la

Tabla 3 se detalla a breves rasgos las historias de usuario generadas para el desarrollo del plugin.

Tabla 3. Historias de usuario para desarrollar el Plugin.

Nro. de Historia	Historia de Usuario	Tareas Identificadas
1	El usuario necesita que el plugin obtenga métricas sobre el impacto de un artículo en las redes sociales, bases de datos bibliográficas y gestores bibliográficos, y posteriormente las muestre de una forma que sea fácil de interpretar.	<ul style="list-style-type: none">• Consumo de APIs de redes sociales y bases de datos de citas.• Mostrar los datos consultados de una manera vistosa y sugerente
2	Se necesita que la información presentada en el artículo no sea tan simple, los datos deben ser más vistosos y diferenciarse los rasgos del artículo mediante colores, títulos, pestañas, entre otras formas, no se debe presentar un artículo con datos mezclados, deben poderse diferenciar los rasgos de los mismos.	<ul style="list-style-type: none">• Agregar estilos CSS (Bootstrap) a las plantillas.• Agrupar la información presentada mediante pestañas.
3	Es necesario que el contenido mostrado dentro de una publicación en OJS se pueda adaptar a cualquier dispositivo tanto de escritorio como móviles.	<ul style="list-style-type: none">• A más de la inclusión de Bootstrap, habrá que transformar las plantillas a HTML5.

Fuente: Autor.

4.3.2. Diseño.

Dentro de la fase de Diseño y mediante las historias de usuario se han identificado los siguientes requisitos:

Tabla 4: tabla de requisitos para Plugin OJS.

Historia de usuario	Requisito
1	Implementar métricas en el artículo que permitan medir el impacto de una publicación dentro de las redes sociales.
1	Mejorar la interfaz y la presentación de las métricas dentro de un artículo publicado en OJS, de forma que sea más sencillo para el usuario interpretar la información obtenida y presentada.
2	Mejorar el diseño y la presentación de la información dentro de los artículos científicos.
2	Agrupar la información mediante pestañas para poder distinguir los diferentes datos del artículo.
3	Adaptar la presentación de los artículos a cualquier dispositivo, mejorando la adaptabilidad no sólo a computadoras portátiles, pantallas de escritorio, sino también a dispositivos móviles como tablets y smartphones.

Fuente: Autor.

También para la fase de diseño se tomó como modelo la presentación de los artículos mostrados en PLOS:

plos.org create account sign in

PLOS BIOLOGY Browse Publish About Search advanced search

OPEN ACCESS FEER-REVIEWED

RESEARCH ARTICLE

Near-Stasis in the Long-Term Diversification of Mesozoic Tetrapods

Roger B. J. Benson, Richard J. Butler, John Alroy, Philip D. Mannion, Matthew T. Carrano, Graeme T. Lloyd

Published: January 25, 2016 • <http://dx.doi.org/10.1371/journal.pbio.1002359>

Article Authors Metrics Comments Related Content

22 Saves 0 Citations

6,091 Views 157 Shares

Download PDF Print Share

CrossMark

Included in the Following Collection

Paleontology

Subject Areas

- Mesozoic era
- Species diversity
- Paleogeography
- Phytogeography
- Fossils
- Speciation
- Biodiversity
- Fossil record

ADVERTISEMENT

PLOS

Meta-Research: Improving Research

PLOS Biology Collection

Submissions welcome

Total Article Views: **6,091**

Invalid date (publication date) through Apr 17, 2016 *

	HTML Page Views	PDF Downloads	XML Downloads	Totals
PLOS	5,477	594	11	6,082
PMC	5	3	n.a.	9
Totals	5,483	597	11	6,091

10.89 % of article views led to PDF downloads

Cumulative Views

Months

*Although we update our data on a daily basis, there may be a 48-hour delay before the most recent numbers are available. PMC data is posted on a monthly basis and will be made available once received.

figshare 92

Cited

DataCite 1

Google Scholar Search

Saved

MENDELEY 22

Discussed

twitter 41

facebook 116

Comments 0

Information on PLOS Article-Level Metrics
Questions or concerns about usage data? Please let us know.

Figura 13: presentación de un artículo en Plos Biology
Fuente: Revista de ejemplo publicada en:
<http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002359>

Como parte de la fase de diseño se programaron reuniones de presentación y análisis con el usuario de los requisitos que se van a implementar, para confirmar si está de acuerdo con los mismos y si se tiene versiones prototipo presentar dichas implementaciones, esto con el fin de discutir modificaciones en los requerimientos según las observaciones del usuario y posteriormente implementarlos luego de analizar la viabilidad de los mismos, estas reuniones dieron como resultado versiones del Plugin.

4.3.2.1. Versiones e Iteraciones.

Las versiones son los programas funcionales según el alcance acordado con el usuario, son prototipos del plugin que se definen según los requisitos del usuario, mientras que las iteraciones son el número de veces que se han realizado cambios sobre un prototipo del plugin, pudiendo estas extenderse por cada versión o prototipo implementado.

- **Versión 0.1**

Iteración 1:

Se realizó un análisis de las posibles maneras de extraer información para luego contabilizarla y crear métricas para mostrarlas al visitante del sitio de la revista en cada artículo. En este análisis se determinó que la manera más óptima y rápida de obtener los datos, era el consumo del API de herramientas que ya hiciesen por si solas el cálculo.

Como el impacto de un artículo se lo puede medir de acuerdo a las veces que ha sido difundido a través de redes sociales tanto de ocio como de redes destinadas a fines científicos; y debido a que una de las herramientas existentes abarca la mayoría de estas redes sociales, se optó por consumir el API de ALTMETRIC. Además se incluyó el API de ELSEVIER el cual brinda datos de citas bibliográficas, otro dato que es relevante para una construcción de métricas.

En esta primera iteración se construyó un script el cual mediante el parámetro DOI, que la mayoría de artículos científicos posee, se realizó la búsqueda de información en las APIs antes mencionadas para recuperar resultados que expresen de una manera numérica la importancia y penetración del artículo en cuestión.

Los datos obtenidos se encontraban en un formato JSON el cual se analizó para filtrar únicamente la información importante que serían tomados para construir el conjunto de métricas a mostrar, siendo las de relevancia las siguientes: la cantidad de veces citado de SCOPUS, la cantidad de veces mencionado en las redes sociales Google+, Facebook,

Twitter y Reddit; y la cantidad de veces que ha sido guardado en las redes CiteULike y Mendeley.

Iteración 2:

Luego de haber filtrado los datos relevantes, se creó una página de ejemplo la cual se diseñó rápidamente y solo para efectos de demostración; esta página dio a conocer de una manera visual y mediante texto plano los datos consultados en tiempo real.

Esta versión se mostró al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados. Se aclaró que esta versión servirá para cumplir con la historia de usuario número 1.

- **Versión 0.2**

Iteración 1:

Se creó el primer prototipo del plugin el cual se puede instalar en OJS, este plugin agrega en la página de visualización del artículo las métricas a manera de datos estadísticos mediante la construcción de un diagrama de barras para mostrar cada dato numérico y el lector del artículo pueda comparar el impacto contrastado entre cada red social.

Esta versión se trabajó bajo los requerimientos de las historias de usuario número 2 y 3.

El cliente luego de ver el avance, solicitó que la información se presente de otra manera y para ello mostró un artículo en la revista PLOS Biology. El cliente también sugirió tomar ideas de la visualización para adaptarla a OJS mediante el plugin.

- **Versión 0.3**

Iteración 1:

Luego de analizar la forma de presentar los datos en PLOS Biology, además de la estructura visual; se modificó las plantillas que OJS provee.

A cada plantilla se la modificó para que estuviera dentro del standard HTML5 y se agregaron las librerías CSS de Bootstrap para cambiar la apariencia. En cada plantilla también se cambiaron los estilos para adaptarlos a un diseño responsivo y el artículo se pueda observar en cualquier dispositivo.

Iteración 2:

Luego de modificar las plantillas para que la apariencia sea congruente con la historia de usuario número 2 y 3, se cambió la forma de presentar los resultados según el cliente lo solicitó. Ahora ni los datos de las redes sociales, ni de SCOPUS, se presentarán a manera de diagrama de barras. En su lugar se cargarán los datos de visitas y descargas de la propia revista en dicho diagrama.

Ahora para presentar los datos extraídos se ha tomado en consideración la forma en la cual PLOS Biology lo realiza, y a cada dato se le agregó la imagen de la red a la cual pertenece.

Esta versión se la presentó al cliente, el cual se mostró conforme con el avance presentado a la vez que solicitó que la distribución de contenidos del artículo también se asemeje al observado en PLOS Biology.

- **Versión 0.4**

Iteración 1:

Con el nuevo requerimiento del cliente, se tuvo que modificar la plantilla de presentación de los artículos para agregar una barra de pestañas las cuales mostraran información referente al artículo observado asemejándose a lo observado en PLOS Biology.

Esta versión se presentó al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados.

En la Figura 14 podemos ver el cambio del menú de OJS original (A) por el nuevo diseño (B) obtenido gracias al uso de BootStrap que nos ayuda a hacer un menú responsivo (C) cuando el tamaño de la pantalla es pequeño.

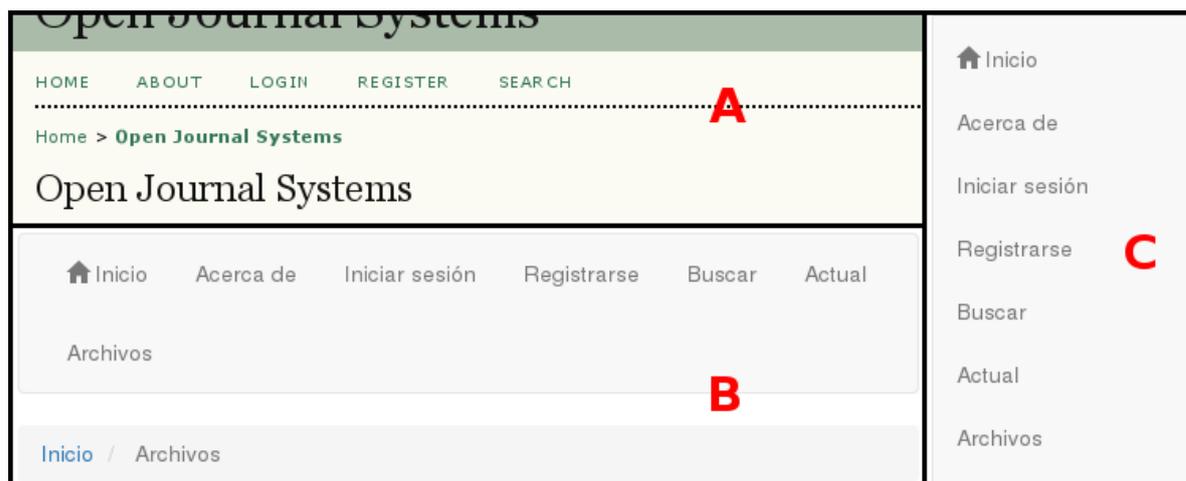


Figura 14: Barra de Navegación de OJS.

Fuente: Revista de ejemplo OJS ubicada en <http://metrics.cu.cc/index.php/utpl>

4.3.3. Desarrollo.

Se puede dividir el desarrollo en 3 puntos:

- ✓ Extracción de métricas.- son el consumo de los servicios web de Scopus y Almetric para obtener las métricas para cada artículo.
- ✓ Procesamiento de datos.- es el procesamiento de las métricas obtenidas.
- ✓ Pantallas y visualización.- son las interfaces diseñadas e implementadas según los requerimientos del usuario.

4.3.3.1. Extracción de Métricas.

Para la extracción de métricas de artículos se necesita primero obtener los DOI de cada publicación, para ello se debe consultar los servicios de SCOPUS y ALTMETRIC, se detalla a continuación la consulta de cada uno de ellos.

4.3.3.2. ELSEVIER.

Dentro del servicio de búsqueda de SCOPUS, mediante su herramienta HTML, existen varias opciones para la búsqueda de datos, se puede solicitar información por atributos generales como: autores, editoriales, títulos de artículos, títulos de revistas y muchos más; pero también están la búsqueda a través de identificadores únicos como el ISSN (International Standard Serial Number; Número Internacional Normalizado de Publicaciones Seriadas), el CODEN (código alfanumérico de seis caracteres) o el DOI (Digital Object Identifier).

Mediante el API de ELSEVIER podemos acceder a las mismas opciones de búsqueda pero con respuestas más específicas en formato XML⁵. Dentro del Anexo D se detalla la búsqueda de citas mediante el API que provee ELSEVIER.

Con la ayuda de dicho manual se puede saber que la sintaxis de la URL para la búsqueda de información en SCOPUS es de la siguiente manera:

API URL + API Key + DOI



Figura 15: URL formada para la consulta del DOI en el API de Altmetric.

Fuente: Autor.

Cuando toda esta cadena de caracteres está bien formada, el navegador nos devuelve la información en formato XML, el cual se puede leer para obtener las etiquetas que sean más relevantes. Para efectos de este capítulo, la única etiqueta que nos interesa es la que hace referencia al conteo de citas que sería: “<citedby-count></citedby-count>”.

```
<?xml version="1.0" encoding="UTF-8" ?>
<search-results xmlns="http://www.w3.org/2005/Atom" xmlns:atom="http://www.w3.org/2005/Atom" xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/" xmlns:prism="http://prismstandard.org/namespaces/basic/1.2/">
  <opensearch:totalResults>1</opensearch:totalResults>
  <opensearch:startIndex>0</opensearch:startIndex>
  <opensearch:itemsPerPage>1</opensearch:itemsPerPage>
  <opensearch:Query role="request" searchTerms="DOI%28%2210.1371%2Fjournal.pone.0069841%22%29apiKey=3&start=0&count=25&query=DOI%28%2210.1371%2Fjournal.pone.0069841%22%29apiKey=3&start=0&count=25&query=DOI%28%2210.1371%2Fjournal.pone.0069841%22%29apiKey=3" />
  <entry>
    <link href="http://api.elsevier.com/content/abstract/scopus_id/84881488287" />
    <link href="http://www.scopus.com/inward/record.url?partnerID=Hz0xMe3b&scopusId=84881488287" />
    <link href="http://www.scopus.com/inward/citedby.url?partnerID=Hz0xMe3b&scopusId=84881488287" />
    <prism:url>http://api.elsevier.com/content/abstract/scopus_id/84881488287</prism:url>
    <dc:identifier>SCOPUS_ID:84881488287</dc:identifier>
    <eid>2-s2.0-84881488287</eid>
    <dc:title>Facebook Use Predicts Declines in Subjective Well-Being in Young Adults</dc:title>
    <dc:creator>Kross, E.</dc:creator>
    <prism:publicationName>PLOS ONE</prism:publicationName>
    <prism:eIssn>19326203</prism:eIssn>
    <prism:volume>8</prism:volume>
    <prism:issueIdentifier>8</prism:issueIdentifier>
    <prism:coverDate>2013-08-14</prism:coverDate>
    <prism:coverDisplayDate>14 August 2013</prism:coverDisplayDate>
    <prism:doi>10.1371/journal.pone.0069841</prism:doi>
    <citedby-count>45</citedby-count>
    <affiliation>
      <affilname>University Michigan Ann Arbor</affilname>
      <affiliation-city>Ann Arbor</affiliation-city>
      <affiliation-country>United States</affiliation-country>
    </affiliation>
    <pubmed-id>23967061</pubmed-id>
    <prism:aggregationType>Journal</prism:aggregationType>
    <subtype>ar</subtype>
    <subtypeDescription>Article</subtypeDescription>
    <article-number>e69841</article-number>
    <source-id>10600153309</source-id>
  </entry>
</search-results>
```

Figura 16: XML de respuesta de SCOPUS a través del navegador Firefox.

Fuente: Autor.

⁵eXtensibleMarkupLanguage (lenguaje de marcas extensible), es un lenguaje de marcas desarrollado por el World Wide Web Consortium (W3C) utilizado para almacenar datos en forma legible.

Para el consumo de la API de ELSEVIER en el presente proyecto específicamente para el consumo del recurso de búsquedas de SCOPUS a nivel de programación, basta con la concatenación de caracteres siguiendo la sintaxis para conformar la URL de búsqueda utilizando el código de la Figura 17:

```
function _getCitas($doi)
{
    $citas=array();
    $scopus_url = ELSEVIER_URL.$this->scopusKey."&query=DOI(\"\".$doi.\"");
    $scopus_response = file_get_contents($scopus_url);
    $scopus_json = json_decode($scopus_response);
    if ($scopus_json->{'search-results'}->{'opensearch:totalResults'}==0){
        $citas+=array("scopus"=>0);
    } else{
        $citas+=array("scopus"=>intval($scopus_json->{'search-results'}->{'entry'}[0]->{'
            citedby-count'}));
    }
    .
    .
    .
    return array($citas,$altmetricMoreInfo);
}
```

Figura 17: Código para obtener información de ELSEVIER.

Fuente: Autor.

4.3.3.2.1. Altmetric.

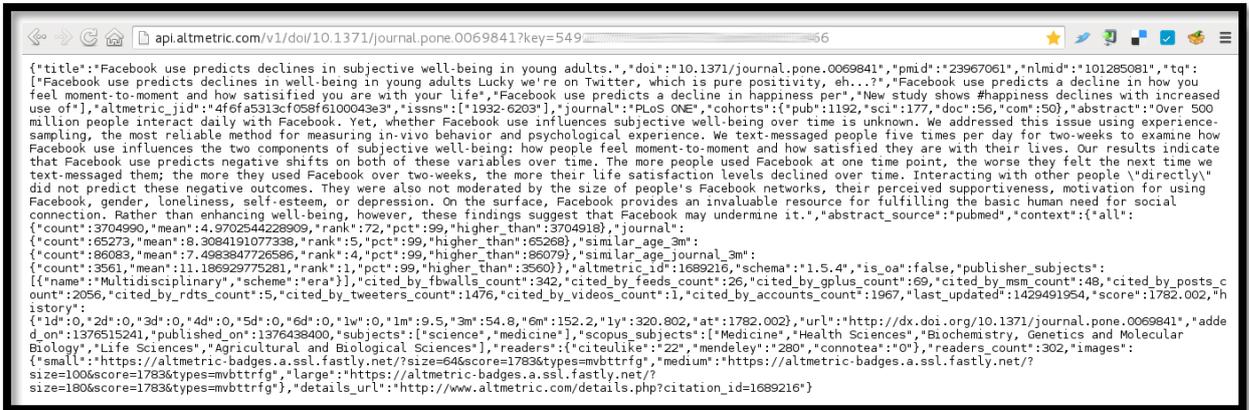
El servicio de Altmetric se lo utiliza para obtener información de varios sitios web, en su mayoría redes sociales, al acceder a una sola URL. Altmetric recopila la información de varias redes sociales tanto de ocio como las consideradas para fines investigativos como son Mendeley o CiteULike.

Similar a ELSEVIER, Altmetric hace uso de una API Key para obtener resultados; esta llave no es obligatoria al momento de realizar la petición de información, pero solo se podrá realizar una cantidad limitada de consultas (720 peticiones de información cada hora), es decir que Altmetric posee una forma menos restrictiva de brindar la información al usuario.

Una vez obtenida la API Key, se prueba en un navegador cómo funciona el API de Altmetric; para ello se debe formar la URL que se parece en algo a la de SCOPUS, pero con algunas variantes de sintaxis: API URL + DOI + API Key.

<http://api.altmetric.com/v1/doi/10.1371/journal.pone.0069841?key=APIKEYOBTENIDO>

Al navegar hacia la URL formada, obtendremos un documento en formato JSON el cual nos dará toda la información que existe para dicho DOI.



```
{
  "title": "Facebook use predicts declines in subjective well-being in young adults.",
  "doi": "10.1371/journal.pone.0069841",
  "pmid": "23967061",
  "nlmid": "101285081",
  "tq": [
    "Facebook use predicts declines in well-being in young adults Lucky we're on Twitter, which is pure positivity, eh...?",
    "Facebook use predicts a decline in how you feel moment-to-moment and how satisfied you are with your life",
    "Facebook use predicts a decline in happiness per",
    "New study shows #happiness declines with increased use of",
    "altmetric_jid": "4f6fa5313cf058f6100043e3",
    "issns": [
      "1932-6203"
    ],
    "journal": "PLOS ONE",
    "cohorts": {
      "pub": 1192,
      "sci": 177,
      "doc": 56,
      "com": 50
    },
    "abstract": "Over 500 million people interact daily with Facebook. Yet, whether Facebook use influences subjective well-being over time is unknown. We addressed this issue using experience-sampling, the most reliable method for measuring in-vivo behavior and psychological experience. We text-messaged people five times per day for two-weeks to examine how Facebook use influences the two components of subjective well-being: how people feel moment-to-moment and how satisfied they are with their lives. Our results indicate that Facebook use predicts negative shifts on both of these variables over time. The more people used Facebook at one time point, the worse they felt the next time we text-messaged them; the more they used Facebook over two-weeks, the more their life satisfaction levels declined over time. Interacting with other people \"directly\" did not predict these negative outcomes. They were also not moderated by the size of people's Facebook networks, their perceived supportiveness, motivation for using Facebook, gender, loneliness, self-esteem, or depression. On the surface, Facebook provides an invaluable resource for fulfilling the basic human need for social connection. Rather than enhancing well-being, however, these findings suggest that Facebook may undermine it.",
    "abstract_source": "pubmed",
    "context": {
      "all": {
        "count": 3704990,
        "mean": 4.9702544228909,
        "rank": 72,
        "pct": 99,
        "higher_than": 3704918
      },
      "journal": {
        "count": 86273,
        "mean": 8.3084191077338,
        "rank": 5,
        "pct": 99,
        "higher_than": 65268
      },
      "similar_age_3m": {
        "count": 86083,
        "mean": 7.4983847726586,
        "rank": 4,
        "pct": 99,
        "higher_than": 86079
      },
      "similar_age_journal_3m": {
        "count": 3561,
        "mean": 11.186929775281,
        "rank": 1,
        "pct": 99,
        "higher_than": 3560
      },
      "altmetric_id": 1689216,
      "schema": "1.5.4",
      "is_oa": false,
      "publisher_subjects": [
        {
          "name": "Multidisciplinary",
          "scheme": "era"
        }
      ],
      "cited_by_fbwalls_count": 342,
      "cited_by_feeds_count": 26,
      "cited_by_gplus_count": 69,
      "cited_by_msm_count": 48,
      "cited_by_posts_count": 2056,
      "cited_by_rds_count": 5,
      "cited_by_tweeters_count": 1476,
      "cited_by_videos_count": 1,
      "cited_by_accounts_count": 1967,
      "last_updated": 1429491954,
      "score": 1782.002,
      "history": {
        "id": 0,
        "2d": 0,
        "3d": 0,
        "4d": 0,
        "5d": 0,
        "6d": 0,
        "1w": 0,
        "1m": 9.5,
        "3m": 54.8,
        "6m": 152.2,
        "1y": 320.802,
        "at": 1782.002,
        "url": "http://dx.doi.org/10.1371/journal.pone.0069841",
        "add_d_on": 1376515241,
        "published_on": 1376438400,
        "subjects": [
          "science",
          "medicine"
        ],
        "scopus_subjects": [
          "Medicine",
          "Health Sciences",
          "Biochemistry, Genetics and Molecular Biology",
          "Life Sciences",
          "Agricultural and Biological Sciences"
        ],
        "readers": {
          "citeulike": 22,
          "mendeley": 280,
          "connotea": 0
        },
        "readers_count": 302,
        "images": {
          "small": "https://altmetric-badges.a.ssl.fastly.net/?size=64&score=1783&types=mbvttfrfg",
          "large": "https://altmetric-badges.a.ssl.fastly.net/?size=100&score=1783&types=mbvttfrfg",
          "medium": "https://altmetric-badges.a.ssl.fastly.net/?size=180&score=1783&types=mbvttfrfg"
        },
        "details_url": "http://www.altmetric.com/details.php?citation_id=1689216"
      }
    }
  }
}
```

Figura 18: Información del DOI mediante el API Altmetric.

Fuente: <http://api.altmetric.com/>

En este sitio podemos observar que las búsquedas se pueden realizar únicamente a través del DOI y que el resultado que el API nos retorna está en formato JSON. Además indica que la sintaxis para el correcto uso del API como se detalla a continuación:



Figura 19: URL formada para la consulta del DOI en el API de Altmetric.

Fuente: Autor.

Una vez que la petición se realiza correctamente, Altmetric retorna la información del DOI buscado en formato JSON el cual nos ofrece seis etiquetas con información relevante para enriquecer la información del artículo cuantitativamente.

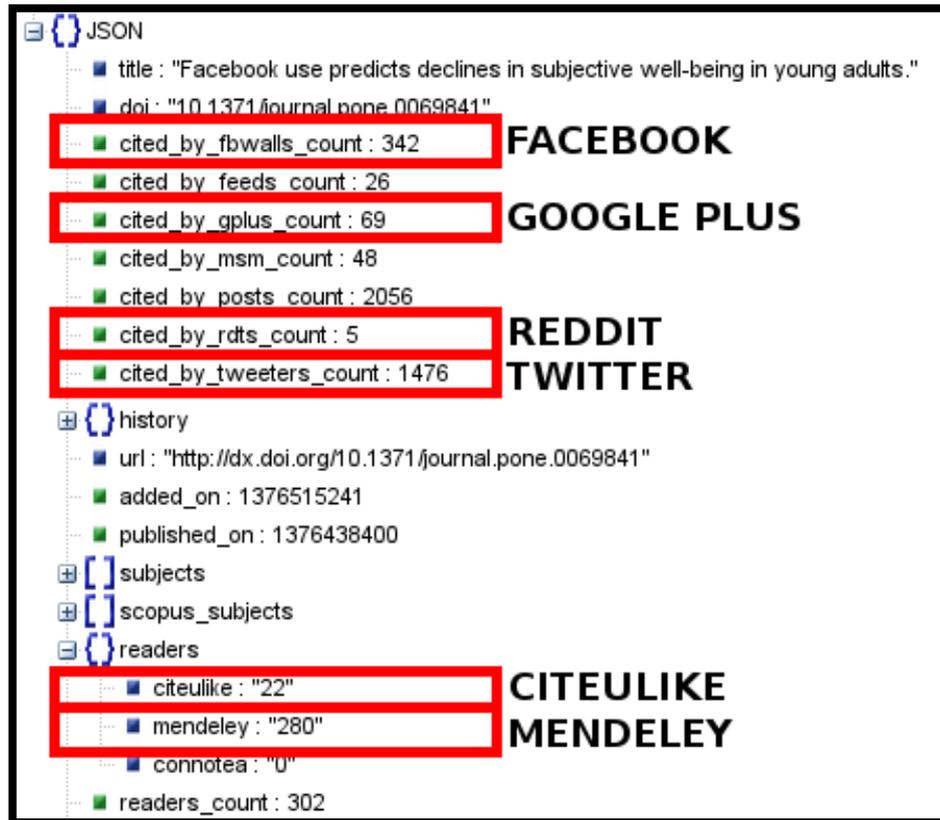


Figura 20: JSON de Altmetric a través de <http://jsonviewer.stack.hu/>
Fuente: Autor.

4.3.3.3. *Procesamiento de datos.*

El proceso de extracción de métricas obtiene los DOI de cada artículo y la información específica del mismo tanto en SCOPUS como en ALTMETRIC, una vez obtenidos todos estos enlaces se procede a almacenarlos en una lista de objetos para su presentación.

Para el procesamiento de esta lista en la página de presentación de un artículo se consumió Chart.js, que es un proyecto de código abierto disponible en Github el cual nos permite mediante JavaScript crear gráficas para representación de datos, esto con el fin de cargar dinámicamente la tabla y las gráficas de las métricas del propio OJS, y los recuadros de información de métricas externas.

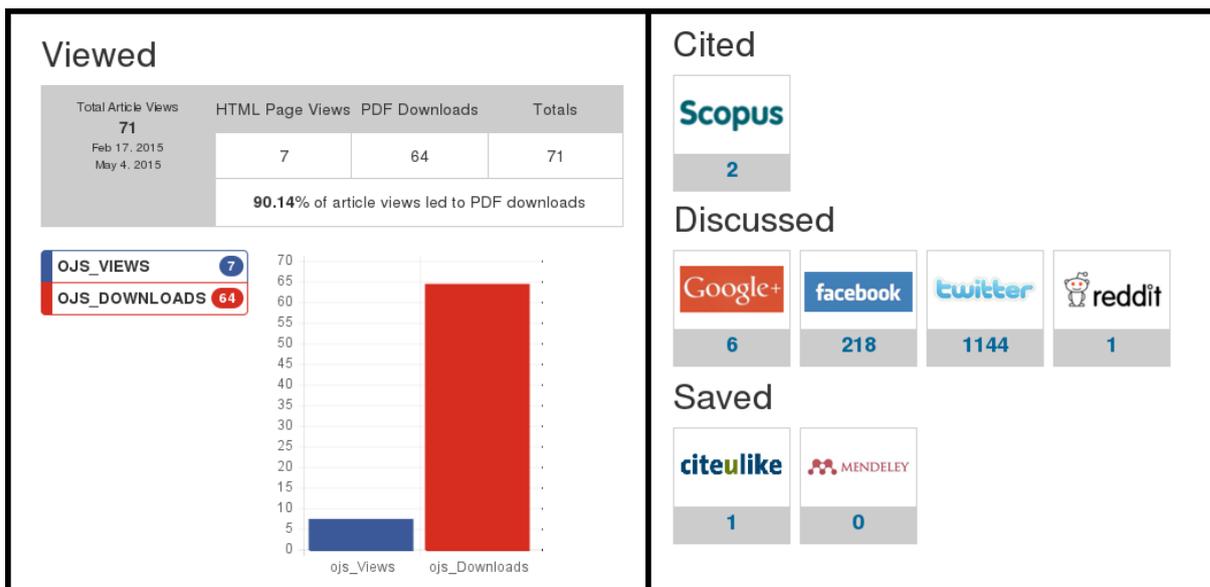


Figura 21: (izq.) Tabla y gráfica de las métricas de OJS. (der.) Recuadro de información externa.
Fuente: Autor. Revista de ejemplo OJS ubicada en <http://metrics.cu.cc/index.php/utpl>

Mediante Chart.js se pueden representar gráficamente datos como por ejemplo el diagrama de barras de la parte izquierda en la Figura 21. De igual manera jQuery, también código abierto, es un proyecto desarrollado con JavaScript que nos ayuda a manejar elementos de HTML de una manera mucho más ágil, esto nos permite por ejemplo crear dinámicamente elementos como los recuadros de información externa de la parte derecha de la Figura 21.

4.3.3.4. Pantallas y visualización.

La implementación de los requisitos del usuario se muestra en las siguientes interfaces:

- a) Página inicial: en esta se presenta la información de forma más vistosa que la presentación anterior, de tal forma que el usuario se pueda enfocar en los datos relevantes o pueda distinguir y encontrar fácilmente la información del artículo, tal como se puede apreciar en la Figura 22 :

The screenshot shows a web interface for a journal article. At the top, a grey header contains the text "NUEVA REVISTA". Below this is a navigation menu with links for Home, About, User Home, Search, Current, and Archives. On the left side, there are buttons for "OPEN JOURNAL SYSTEMS" and "Journal Help". The main content area displays the article title "Instrumentational Complexity of Music Genres and Why Simplicity Sells" by "Stefan Thurner". Below the title are buttons for "Cite Article" and "Share", and a DOI link: "http://dx.doi.org/10.1371/journal.pone.0115255". The article is categorized under "Abstract", "Article Metrics", "About Authors", and "Comments". The abstract text discusses listening habits and music complexity. On the right side, there are several utility boxes: "User You are logged in as... admin" with links for "My Journals", "My Profile", and "Log Out"; "Notifications" with links for "View (25 new)" and "Manage"; "Language" with a dropdown menu set to "English" and a "Submit" button; "Journal Content Search" with a search bar and a "Search" button; "Browse" with links for "By Issue", "By Author", "By Title", and "Other Journals"; and "Font Size" with buttons for "A", "A", and "A".

Figura 22: Artículo de la revista de ejemplo. Pestaña principal Abstract.

Fuente: Autor. Revista de ejemplo ubicada en <http://metrics.cu.cc/index.php/utpl>

b) Pestañas: se ha dividido la información en Pestañas de acuerdo a los datos que se van a presentar

- “**Article Metrics**” muestra información relevante del artículo gracias a los API de SCOPUS y Altmetric, así como también del impacto que ha tenido en artículo en el mismo OJS. Esto aporta ayuda visual para que el lector se dé una idea de cuán relevante es la información que está consumiendo.

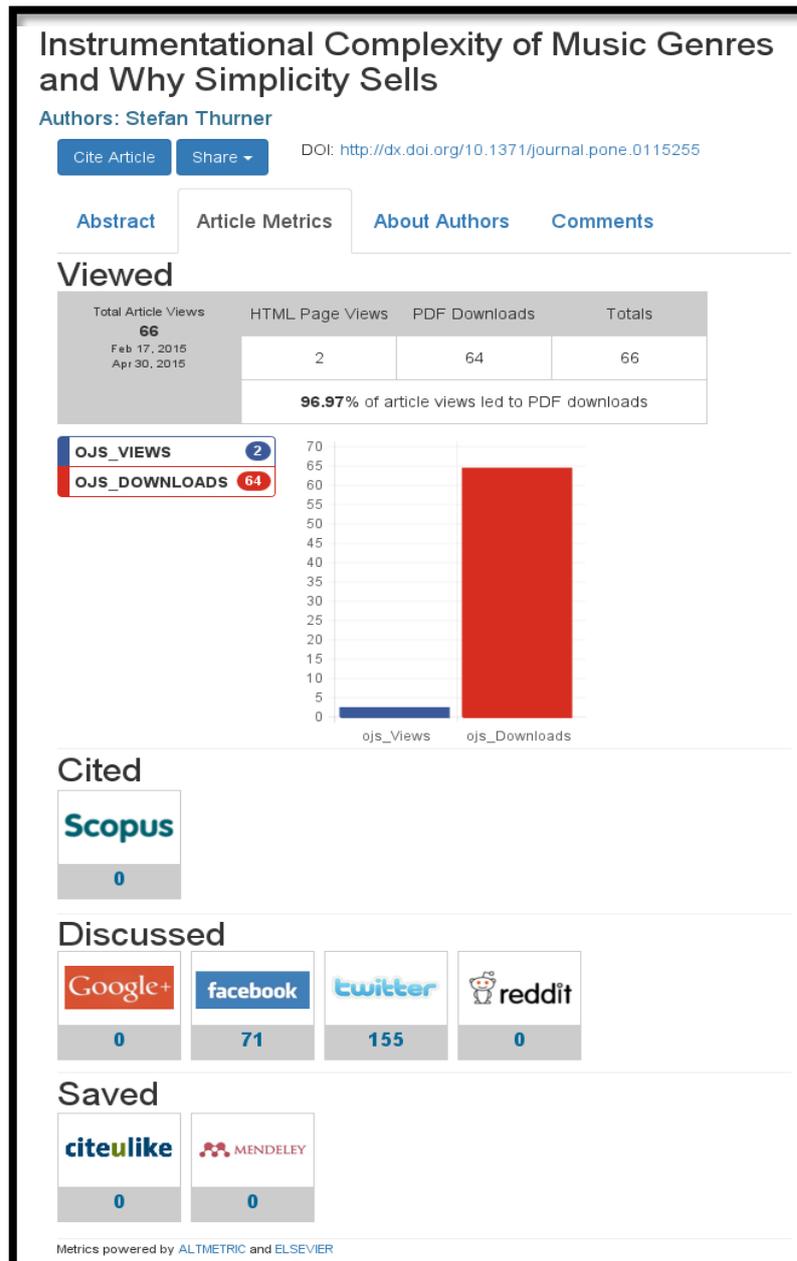


Figura 23: Pestaña métricas de la revista de ejemplo.
Fuente: Autor. <http://metrics.cu.cc/index.php/utpl>

- **About Authors** detalla la biografía y el contacto de cada autor del artículo como vemos en el ejemplo de la Figura 24:



The image shows a screenshot of a journal article's 'About Authors' page. At the top, the article title is 'Instrumentational Complexity of Music Genres and Why Simplicity Sells'. Below the title, the author's name 'Stefan Thurner' is listed. There are buttons for 'Cite Article' and 'Share', and a DOI link: <http://dx.doi.org/10.1371/journal.pone.0115255>. A navigation bar contains links for 'Abstract', 'Article Metrics', 'About Authors' (which is highlighted), and 'Comments'. The main section is titled 'Thurner, Stefan' and contains a 'Biografía:' section with a detailed paragraph about his academic and professional background. At the bottom, there is a 'Contact:' section with the email address stefan.thurner@meduniwien.ac.at.

Figura 24: Pestaña autores de la revista de ejemplo.

Fuente: Autor. Revista de ejemplo OJS ubicada en <http://metrics.cu.cc/index.php/utpl>

- **Comments** en esta pestaña los usuarios pueden encontrar información dejada por otros usuarios sobre el contenido de un artículo, la utilidad del mismo, retroalimentación o artículos relacionados, puesto que los comentarios son una gran ayuda al momento de juzgar el contenido de un artículo, debido a ello se los ha colocado en una pestaña más accesible como se muestra en la Figura 25.



Figura 25: Pestaña comentarios de la revista de ejemplo.

Fuente: Autor. Revista de ejemplo OJS ubicada en <http://metrics.cu.cc/index.php/utpl>

c) Botones: se han implementado 2 botones que ayuden a la difusión de un artículo :

- **“Cite Article”**, ahora se encuentra accesible fácilmente debido a que la difusión de un artículo es importante para darle la trascendencia que se merece. OJS nos ofrece varios formatos para citar el artículo.

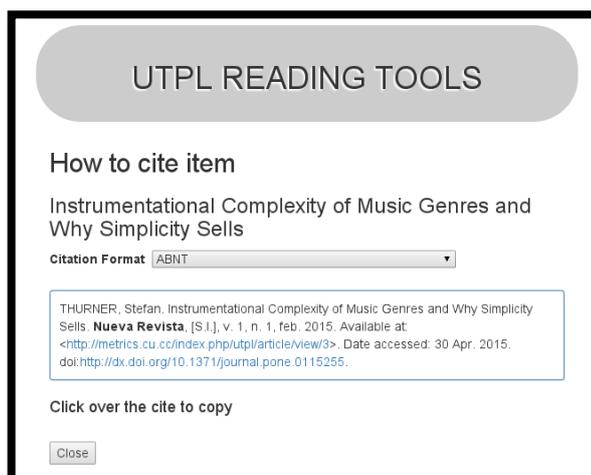


Figura 26: Ventana emergente para elegir la forma de citar el artículo.

Fuente: Autor. Revista de ejemplo OJS ubicada en <http://metrics.cu.cc/index.php/utpl>

Una vez se ha elegido el formato de citación que se quiere utilizar, basta con hacer click sobre la cita construida de acuerdo al formato y una ventana emergente ayudará a copiar directamente toda la cita para evitar errores de pérdida de caracteres al seleccionar manualmente.

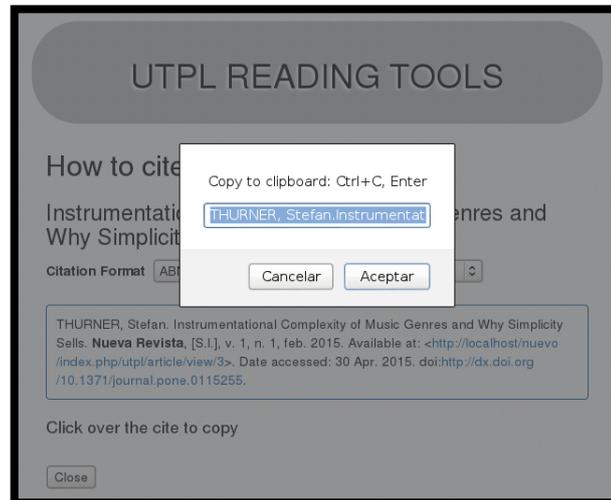


Figura 27: Acción de copia de la cita seleccionada.

Fuente: Autor. Revista de ejemplo OJS ubicada en <http://metrics.cu.cc/index.php/utpl>

- **“Share”** también nos ayuda a difundir la información del artículo pero no a través de una manera convencional, sino enfocado a estas métricas alternativas.



Figura 28: Opción de compartir artículo.

Fuente: Autor. Revista de ejemplo OJS ubicada en <http://metrics.cu.cc/index.php/utpl>

4.3.4. Pruebas.

Para las pruebas realizadas se tomaron en cuenta las historias de usuario para verificar las implementaciones realizadas:

Tabla 5: Pruebas realizadas.

Pruebas Funcionales	Requisito	Resultado
Obtener las métricas sobre el artículo publicado en el OJS	Implementar métricas en el artículo que permitan medir el impacto de una publicación dentro de las redes sociales.	Se visualizan las métricas
Visualizar la correcta presentación de las métricas sobre un artículo en OJS	Mejorar la interfaz y la presentación de las métricas dentro de un artículo publicado en OJS, de forma que sea más sencillo para el usuario interpretar la información obtenida y presentada.	Se muestran gráficos e indicadores con el valor de las métricas obtenidas
Verificar que las métricas obtenidas coincidan con las mostradas en las páginas de Scopus y Almetrics.	Mejorar la interfaz y la presentación de las métricas dentro de un artículo publicado en OJS, de forma que sea más sencillo para el usuario interpretar la información obtenida y presentada.	Las métricas presentadas en OJS coinciden con las mostradas en Scopus y Almetrics
Información se presenta de forma correcta	Mejorar el diseño y la presentación de la información dentro de los artículos científicos.	Los datos se muestran de forma vistosa y fluida, no existe sobre posición de objetos y la interfaz es amigable con el usuario.
Datos se presenten completos.	Mejorar el diseño y la presentación de la información dentro de los artículos científicos.	No existen cadenas de caracteres cortadas y la información se presente de forma completa.
Visualizar que los datos estén correctamente distribuidos en las pestañas.	Agrupar la información mediante pestañas para poder distinguir los diferentes datos del artículo.	No existe incorrecta ubicación de la información, los datos se presentan correctamente.
Visualizar el artículo	Adaptar la presentación de los	Se muestra la información

publicado en OJS en distintas pantallas	artículos a cualquier dispositivo, mejorando la adaptabilidad no sólo a computadoras portátiles, pantallas de escritorio, sino también a dispositivos móviles como tablets y smartphones.	correctamente distribuida y acorde al tamaño de la pantalla en que se visualiza.
--	---	--

Fuente: Autor.

4.4. Desarrollo de la Herramienta de Marketing.

Para el desarrollo e implementación de la herramienta de Marketing se siguen las siguientes fases:

- a) Planificación.- se toman las historias de usuario para definir los requisitos y el modelado del sistema.
- b) Diseño.- en esta fase se identifican los requisitos y se mantienen reuniones con el usuario para mostrar las implementaciones realizadas, esto con el fin de obtener retroalimentación por parte del usuario para implementar nuevos requisitos o modificar los ya existentes.
- c) Desarrollo.- en esta fase se implementan los requisitos en código fuente.
- d) Pruebas.- se establecen pruebas de funcionalidad para verificar el correcto funcionamiento de los requisitos implementados y si satisfacen los requerimientos del usuario.

4.4.1. Planificación.

4.4.1.1. Historias de usuario

Descripciones claras por parte del cliente, sin entrar en detalles de la manera de codificar.

Tabla 6: Historias de usuario para Herramienta de Marketing.

Nro. de Historia	Historia de Usuario	Tareas Identificadas
1	La aplicación debe ser intuitiva y de fácil uso.	<ul style="list-style-type: none">• Diseño de una interfaz intuitiva para el usuario.
2	Obtener un nombre de autor, su correo y el título del artículo para cada referencia del artículo principal, solo si esta referencia posee un DOI y una dirección de correo.	<ul style="list-style-type: none">• Realizar el análisis de la estructura HTML de SCOPUS.• Obtener las referencias que poseen DOI para analizarlas.• Identificar si el artículo referenciado posee un email.
3	Crear un archivo CSV con la información obtenida.	<ul style="list-style-type: none">• Crear una opción de exportar un archivo con la información obtenida en formato CSV.
4	Conservar la información para que en futuras búsquedas sea un proceso más rápido	<ul style="list-style-type: none">• Almacenar los resultados en una base de datos
5	No todos podrán tener acceso a la aplicación ya que los correos de los autores son información personal y solo de contacto.	<ul style="list-style-type: none">• Agregar un módulo de autenticación.• Crear un módulo para el cambio de la contraseña del usuario.

Fuente: Autor.

4.4.2. Diseño.

Dentro de la fase de Diseño y mediante las historias de usuario se han identificado los siguientes requisitos:

Tabla 7: tabla de requisitos para Herramienta de Marketing.

Historia de usuario	Requisito
1	Interfaz amigable para el usuario.
2	Obtener autor, correo y título del artículo
3	Agregar la opción de exportar en un archivo CSV la información obtenida mediante el proceso de extracción.
4	Almacenar las búsquedas en una base de datos para un acceso más rápido a la información solicitada.
5	Crear un módulo de autenticación para la herramienta.

Fuente: Autor.

Al igual que en el Plugin OJS se programaron reuniones con el usuario para discutir los requerimientos que se van a implementar, mostrar versiones prototipo y obtener retroalimentación sobre nuevos requisitos a implementar o actualizaciones en los existentes, esto dio como resultado versiones de la herramienta.

4.4.2.1. Versiones e Iteraciones.

- **Versión 0.1**

Iteración 1:

Se realizó un análisis de la estructura HTML de las búsquedas realizadas manualmente en SCOPUS. Este análisis dio a conocer que el sitio web carga en su mayor parte el contenido con JavaScript lo cual dificulta la extracción de datos ya que solo se obtiene el código HTML parcial de la página.

Con esta búsqueda manual y analizando el código HTML de la página se pudo determinar qué etiquetas nos guiarán a conseguir las referencias del artículo principal para luego

descartar aquellas que no poseen DOI. Además, ya que la estructura HTML es la misma para todos los artículos, se pudo identificar la ruta de etiquetas que permitirían saber si un artículo posee correo electrónico y también la ruta de etiquetas del autor al que pertenece el correo.

Iteración 2:

Luego de encontrar la primera dificultad en la iteración 1, se realizó una búsqueda para encontrar una posible solución; gracias a esta búsqueda se obtuvo la librería llamada “selenium” la cual ayuda a obtener todo el código HTML luego que el código JavaScript ha sido ejecutado. Para realizar esto la librería “selenium” debe apoyarse con un programa, que simula un navegador web, llamado “Phantom.js” ya que según la búsqueda realizada es la mejor combinación para este tipo de trabajos.

En esta iteración de la versión 1, se creó un script funcional el cual extraía correctamente los DOI de las referencias y luego extraía de cada referencia el autor, email y título del artículo solo si existía un email.

Con esta versión se dio solución a la historia de usuario número 2.

Esta versión se le enseñó al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados.

- **Versión 0.2**

Iteración 1:

Se realizó la creación de una base de datos con las siguientes tablas:

- ❖ **Tabla ARTICULOS:** Esta tabla existe con el propósito de almacenar los datos relacionados con el artículo principal al cual se lo analizará para obtener las referencias.
 - ✓ **doi_m:** DOI del artículo.
 - ✓ **autor:** Nombre del autor principal del artículo.
 - ✓ **titulo:** Título del artículo.
- ❖ **Tabla REFERENCIAS:** Esta tabla tiene el propósito de almacenar la información de los artículos que se encuentran referenciados en el artículo principal. Se guardará

información solo si posee un email. La relación entre ARTICULOS y REFERENCIAS es de 1 a N, porque un artículo posee muchas referencias.

- ✓ **doi_m:** DOI del artículo principal.
- ✓ **doi_s:** DOI del artículo referenciado.
- ✓ **autor:** Nombre del autor del artículo referenciado.
- ✓ **titulo:** Título del artículo referenciado.
- ✓ **correo:** Correo electrónico del autor del artículo referenciado.

En esta iteración de la versión se agregó al script creado en la versión 1, un script que almacenará la información en las tablas antes descritas para cumplir con la historia de usuario número 4.

Esta versión se le enseñó al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados.

- **Versión 0.3**

Iteración 1:

Se desarrolló una interfaz gráfica para que el usuario pueda trabajar de una manera adecuada con los scripts creados. Ahora el usuario puede ingresar un DOI para buscar sus referencias y almacenar todos los datos extraídos en la base de datos creada para el efecto.

Esta versión se le enseñó al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados. Además solicitó que se le añada a la interfaz gráfica de usuario la opción de cargar un archivo de texto con una lista de DOI.

- **Versión 0.4**

Iteración 1:

Con el nuevo requerimiento del usuario, en esta iteración se realizó el diseño de la nueva interfaz gráfica que permitirá al usuario optar por ingresar un solo DOI para su análisis o elegir un archivo de texto que contenga una lista de DOI.

Esta versión se le enseñó al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados. Además solicitó que se añada a la interfaz gráfica un panel en el cual se observe de alguna manera el avance de la aplicación.

- **Versión 0.5**

Iteración 1:

Con el nuevo requerimiento del usuario, en esta iteración se agregó un panel a la interfaz gráfica en el cual se muestre texto que indique al usuario el avance del programa y que permita saber cuándo ha finalizado la ejecución.

Esta versión se le enseñó al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados.

- **Versión 0.6**

Iteración 1:

Se desarrolló e implementó el módulo de autenticación para cumplir con la historia de usuario número 5, este módulo en un principio creará un usuario único llamado “admin” con una contraseña prefijada con el fin de que no cualquiera pueda ingresar a la aplicación para obtener los correos de las referencias.

Iteración 2:

Se desarrolló e implementó el módulo para el cambio de contraseña del usuario.

Esta versión se le enseñó al cliente, el cual dio el visto bueno al sentirse conforme con los avances presentados.

4.4.3. Desarrollo

Se puede dividir el desarrollo en los siguientes puntos:

- ✓ Análisis de la estructura HTML.- se necesita conocer la estructura del artículo para saber qué información se debe extraer.
- ✓ Pantallas y visualización.- las distintas interfaces en las que se presenta la información.

- ✓ Procesamiento de datos.- la forma en que se obtienen los correos a partir de los DOI de los artículos.

4.4.3.1. Análisis de la estructura HTML del resumen del artículo.

El objetivo del análisis manual de la estructura HTML en WebScraping es encontrar particularidades para programar un patrón único que no se pueda repetir dentro del mismo código HTML y así obtener los datos dentro de una página web; este proceso se lleva a cabo mediante la identificación de etiquetas y sus distintos atributos.

Como se puede observar en la Figura 29 el resumen de un artículo es presentado mediante el navegador, mientras que en la Figura 30 se muestra el código HTML interpretado por el navegador. Para el Web Scraping es importante identificar las etiquetas html que contienen la información a extraer, en este caso el título, autor y correo se encuentran al principio del resumen, y esto facilita la búsqueda manual de un patrón para posteriormente programar la extracción automática.

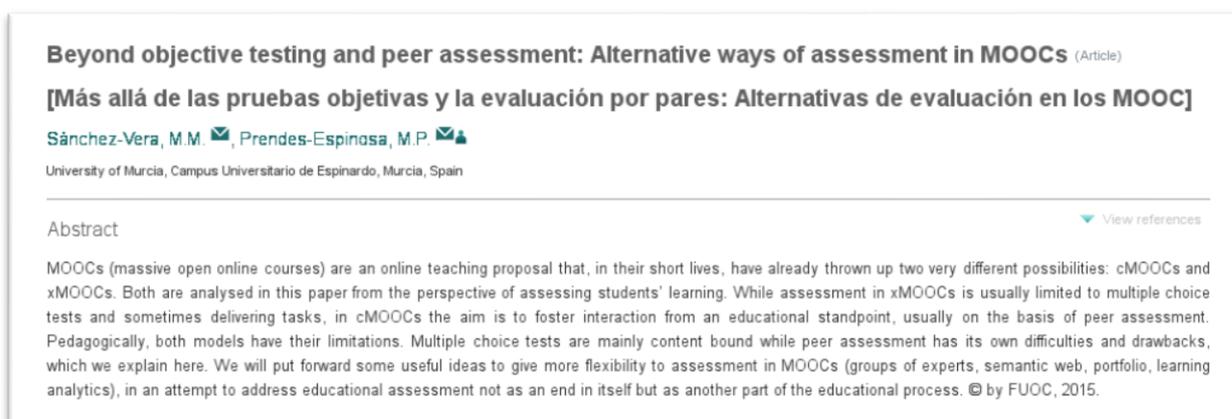


Figura 29: Ejemplo del resumen de un artículo
Fuente: Web de SCOPUS.



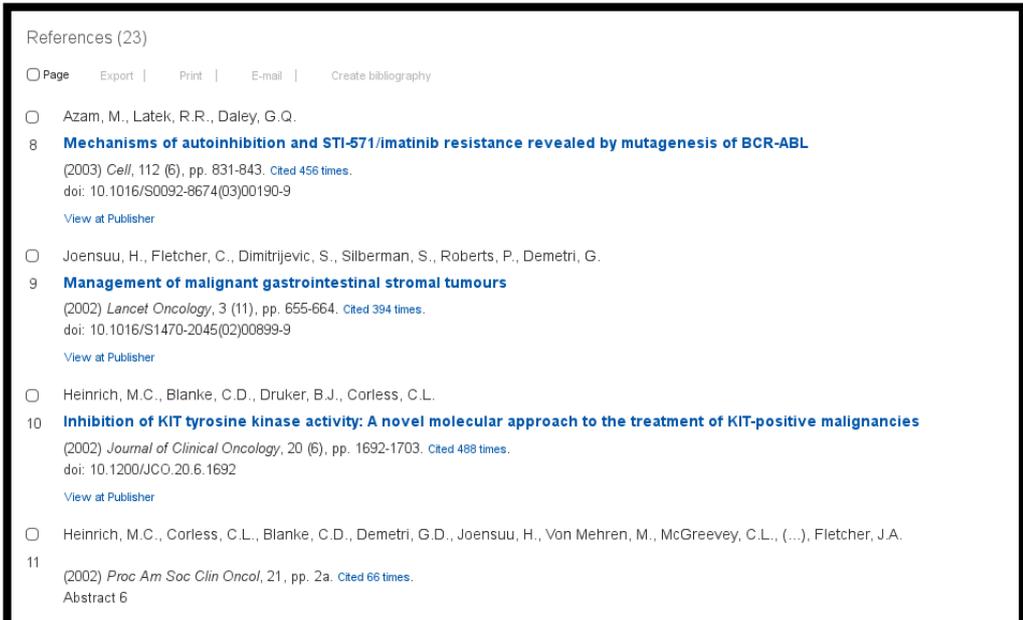
Figura 30: Ejemplo HTML del resumen de un artículo
Fuente: Web de SCOPUS

En el código HTML (Figura 30) se puede observar que el título del artículo se encuentra dentro de una etiqueta “h1” que es la etiqueta estandarizada para colocar el título principal de una página web, mientras que los nombres de los autores del artículo se encuentran cada uno encerrado por las etiquetas “div” y “a” que son etiquetas comúnmente utilizadas para agrupar y crear enlaces a otras páginas respectivamente.

Una vez encontradas las etiquetas generales por decirlo así, debemos hacer que el patrón sea más específico utilizando los atributos “class” de las etiquetas. Mediante esta etiqueta se puede identificar que el correo electrónico de un autor tiene el valor “correspondenceEmail” para dicha etiqueta, que en la Figura 29 está representado por la imagen de un sobre y que está agrupado dentro de una etiqueta “div” con el nombre del autor al que el correo pertenece; también se ha establecido que el título del artículo posee el valor “txtTitle” para el atributo class de la etiqueta “h1”.

De la misma manera debemos analizar el código HTML de las referencias del artículo (Figura 31) para extraer una lista de los artículos referidos que poseen DOI para obtener de estos la base de correos objetivo.

La identificación del patrón es un poco más fácil ya que se va a extraer una lista que en la página de SCOPUS está agrupada dentro de una etiqueta HTML “div” y cada ítem de la lista de referencias dentro de otra etiqueta “div” con el valor “referencesBlk” para el atributo “class”.



References (23)

Page [Export](#) | [Print](#) | [E-mail](#) | [Create bibliography](#)

Azam, M., Latek, R.R., Daley, G.Q.
8 **Mechanisms of autoinhibition and STI-571/Imatinib resistance revealed by mutagenesis of BCR-ABL**
(2003) *Cell*, 112 (6), pp. 831-843. Cited 456 times.
doi: 10.1016/S0092-8674(03)00190-9
[View at Publisher](#)

Joensuu, H., Fletcher, C., Dimitrijevic, S., Silberman, S., Roberts, P., Demetri, G.
9 **Management of malignant gastrointestinal stromal tumours**
(2002) *Lancet Oncology*, 3 (11), pp. 655-664. Cited 394 times.
doi: 10.1016/S1470-2045(02)00899-9
[View at Publisher](#)

Heinrich, M.C., Blanke, C.D., Druker, B.J., Corless, C.L.
10 **Inhibition of KIT tyrosine kinase activity: A novel molecular approach to the treatment of KIT-positive malignancies**
(2002) *Journal of Clinical Oncology*, 20 (6), pp. 1692-1703. Cited 488 times.
doi: 10.1200/JCO.20.6.1692
[View at Publisher](#)

Heinrich, M.C., Corless, C.L., Blanke, C.D., Demetri, G.D., Joensuu, H., Von Mehren, M., McGreevey, C.L., (...), Fletcher, J.A.
11 (2002) *Proc Am Soc Clin Oncol*, 21, pp. 2a. Cited 66 times.
Abstract 6

Figura 31: Referencias de un artículo en SCOPUS

Fuente: Web de SCOPUS

4.4.3.2. Pantallas y visualización.

- a) Pantalla de Autenticación: para el ingreso a la herramienta se presentará primero una interfaz que necesitará de la autenticación a través de un usuario y contraseña, tal como se muestra en la Figura 32.

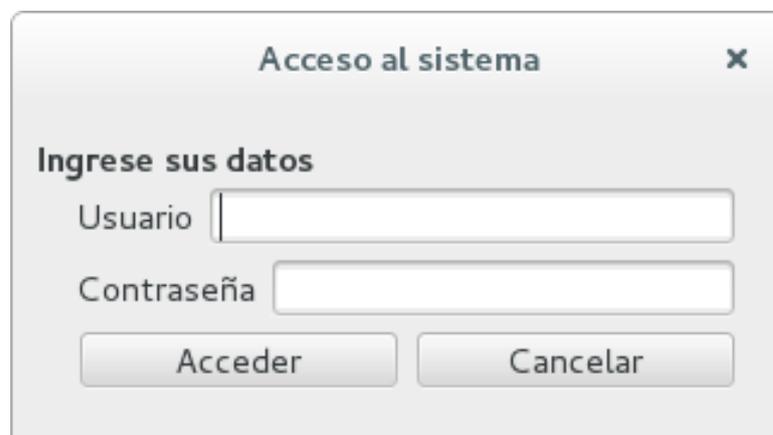
Una ventana de diálogo con el título "Acceso al sistema" y un botón de cerrar "x" en la esquina superior derecha. El contenido principal es "Ingrese sus datos". Hay dos campos de texto: "Usuario" y "Contraseña". Debajo de los campos hay dos botones: "Acceder" y "Cancelar".

Figura 32: ventana de ingreso a la herramienta de marketing.
Fuente: Autor.

Si los datos ingresados son correctos la aplicación accederá a la siguiente interfaz, que es la recuperación de mails.

- b) Pantalla de Búsqueda: interfaz mediante la cual el usuario podrá enviar a recuperar los correos de un artículo o de una lista de artículos mediante su DOI. En la Figura 33 podemos observar en la izquierda la interfaz para ingresar un solo DOI a procesar o en la parte derecha la interfaz para añadir un archivo de texto con la lista de DOI de la cual deseamos obtener los correos.

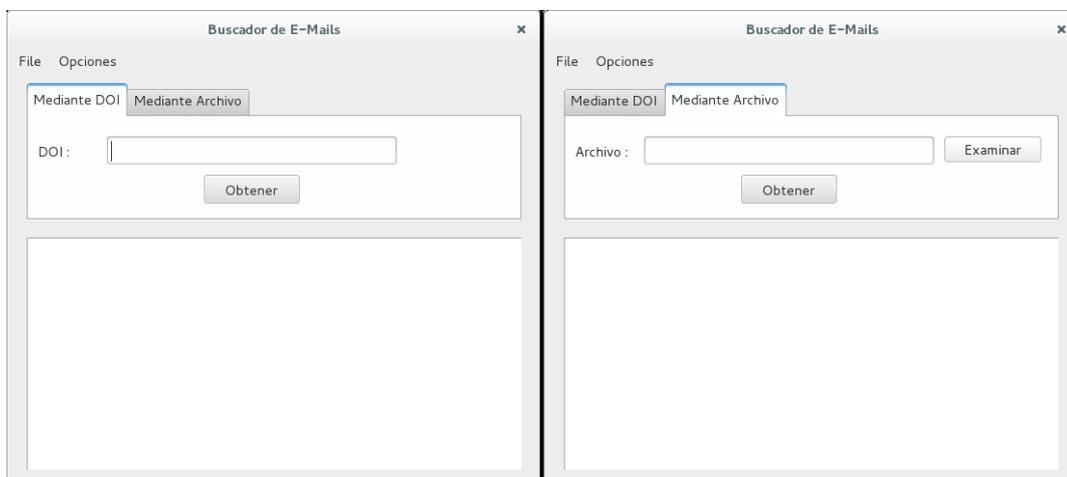
Dos ventanas de diálogo lado a lado, ambas tituladas "Buscador de E-Mails" con un botón de cerrar "x" en la esquina superior derecha. Cada ventana tiene un menú "File" y "Opciones". La ventana de la izquierda tiene pestañas "Mediante DOI" (seleccionada) y "Mediante Archivo". Tiene un campo "DOI:" y un botón "Obtener". La ventana de la derecha tiene pestañas "Mediante DOI" y "Mediante Archivo" (seleccionada). Tiene un campo "Archivo:" con un botón "Examinar" a su derecha y un botón "Obtener" debajo.

Figura 33: formas de ingresar DOI para procesar.
Fuente: Autor.

4.4.3.3. Procesamiento de datos.

La Figura 34 muestra un ejemplo de un archivo de texto que contiene una lista de 13 DOI con la cual el programa puede trabajar para recuperar los correos de las referencias por cada artículo.

```
10.7238/rusc.v12i1.2031
10.7238/rusc.v12i1.1947
10.7238/rusc.v12i1.2078
10.7238/rusc.v12i1.1944
10.7238/rusc.v12i1.1887
10.7238/rusc.v12i1.2475
10.7238/rusc.v12i1.2253
10.7238/rusc.v12i1.2185
10.7238/rusc.v12i1.2260
10.7238/rusc.v12i1.2262
10.7238/rusc.v12i1.2258
10.7238/rusc.v12i1.2296
10.7238/rusc.v12i1.2446
```

Figura 34: Ejemplo del archivo con lista de DOI.

Fuente: Autor.

Luego de introducir ya sea el DOI o el Archivo con los DOI, se realiza el proceso de extracción de datos, el mismo que se conecta con el recurso de ELSEVIER que contiene los resúmenes de los artículos, le envía el DOI del que se quiere analizar y recuperar el nombre del artículo, el nombre del autor y el correo del autor, y posteriormente retorna un XML (Figura 35) que es procesado para obtener la dirección web del resumen del artículo y mediante el cual se adquiere el código HTML completo de la página.

```
<abstracts-retrieval-response>
  <coredata>
    <prism:url>
      http://api.elsevier.com/content/abstract/scopus_id/84921806556
    </prism:url>
    <dc:identifier>SCOPUS_ID:84921806556</dc:identifier>
    <eid>2-s2.0-84921806556</eid>
    <link rel="self" href="http://api.elsevier.com/content/abstract/scopus_id/84921806556"/>
    <link rel="scopus" href="http://www.scopus.com/inward/record.url?partnerID=HzOxMe3b&scp=84921806556&origin=inward"/>
  </coredata>
</abstracts-retrieval-response>
```

Figura 35: XML de respuesta.

Fuente: web de ELSEVIER

Una vez que la herramienta recupera los correos del DOI ingresado o de la lista a través de un archivo, se presenta la información parcial de los correos tal como se muestra en la Figura 36, permitiendo exportar el resultado de la extracción a un archivo CSV.

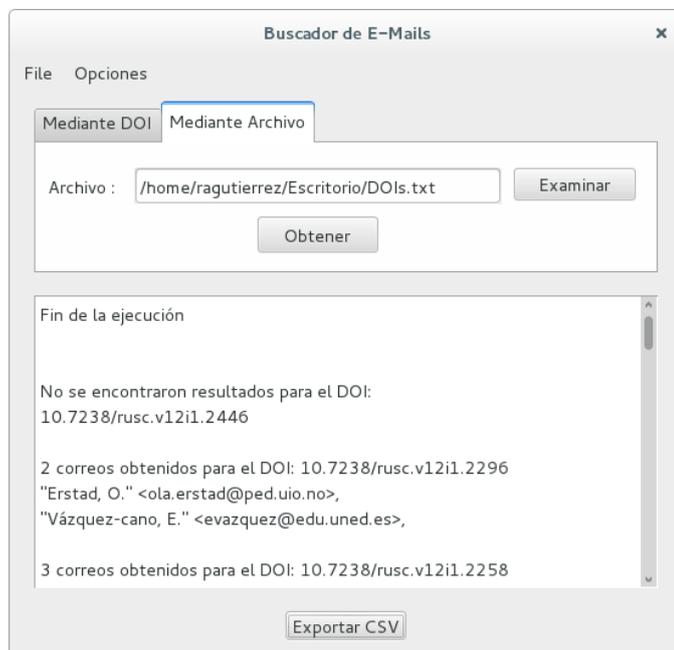


Figura 36: Fin de la ejecución de la herramienta de marketing.
Fuente: Autor.

4.4.4. Pruebas.

4.4.4.1. Pruebas Funcionales.

Se ha realizado una serie de pruebas para comprobar el correcto funcionamiento del programa. En los tres sistemas operativos más utilizados que son: Mac OS (Tabla 8 y Tabla 9), Linux Distribución Debian7 (Tabla 10 y Tabla 11) y Windows 7 (Tabla 12 y Tabla 13).

Para cada una de las pruebas, se ha ejecutado el programa 2 veces con la finalidad de obtener el tiempo de respuesta sin datos almacenados y compararlo con el tiempo de respuesta con datos almacenados.

Tabla 8. Prueba número 1 ejecutada en OSX

DOI Analizado	Nro de Referencias		Nro de correos obtenidos	
	Total	Con DOI	Automáticamente	Manualmente
10.7238/rusc.v11i3.1781	17	2	2	2
10.7238/rusc.v11i3.1778	36	0	0	0
10.7238/rusc.v11i3.1729	29	0	0	0
10.7238/rusc.v11i3.2167	38	6	5	6
10.7238/rusc.v11i3.1808	31	5	3	3
10.7238/rusc.v11i3.2220	38	7	7	7
10.7238/rusc.v11i3.2121	28	3	3	3
10.7238/rusc.v11i3.2090	35	7	4	5
10.7238/rusc.v11i3.2076	43	6	5	5
10.7238/rusc.v11i3.2131	26	5	4	4
10.7238/rusc.v11i3.2102	41	5	5	5

Fuente: Autor.

En la Tabla 8 y Tabla 9 se ejecutó el programa en el sistema operativo Mac OS X, obteniendo un tiempo de ejecución para una lista con 11 DOI de 18 minutos 38 segundos para la primera ejecución y 10 minutos 38 segundos para la segunda ejecución.

Tabla 9. Prueba número 2 ejecutada en OSX

DOI Analizado	Nro de Referencias		Nro de correos obtenidos	
	Total	Con DOI	Automáticamente	Manualmente
10.7238/rusc.v11i3.1781	17	2	2	2
10.7238/rusc.v11i3.1778	36	0	0	0
10.7238/rusc.v11i3.1729	29	0	0	0
10.7238/rusc.v11i3.2167	38	6	6	6
10.7238/rusc.v11i3.1808	31	5	3	3
10.7238/rusc.v11i3.2220	38	7	7	7
10.7238/rusc.v11i3.2121	28	3	3	3
10.7238/rusc.v11i3.2090	35	7	4	5
10.7238/rusc.v11i3.2076	43	6	5	5
10.7238/rusc.v11i3.2131	26	5	4	4
10.7238/rusc.v11i3.2102	41	5	5	5

Fuente: Autor.

En la Tabla 10 y Tabla 11 se ejecutó el programa en el sistema operativo Linux distribución Debian 7, obteniendo un tiempo de ejecución para una lista con 13 DOI de 6 minutos 14 segundos para la primera ejecución y 3 minutos 57 segundos para la segunda ejecución.

Tabla 10. Prueba número 1 ejecutada en Linux distribución Debian7

DOI Analizado	Nro de Referencias		Nro de correos obtenidos	
	Total	Con DOI	Automáticamente	Manualmente
10.7238/rusc.v12i1.2031	42	5	3	3
10.7238/rusc.v12i1.1947	29	0	0	0
10.7238/rusc.v12i1.2078	46	12	12	12
10.7238/rusc.v12i1.1944	37	6	0	6
10.7238/rusc.v12i1.1887	31	11	10	10
10.7238/rusc.v12i1.2475	32	2	1	1
10.7238/rusc.v12i1.2253	26	3	3	3
10.7238/rusc.v12i1.2185	31	5	4	4
10.7238/rusc.v12i1.2260	44	5	5	5
10.7238/rusc.v12i1.2262	34	0	0	0
10.7238/rusc.v12i1.2258	19	3	3	3
10.7238/rusc.v12i1.2296	39	3	2	2
10.7238/rusc.v12i1.2446	0	0	0	0

Fuente: Autor.

En la ejecución de esta lista se observó la particularidad de no haber obtenido resultados con el programa para el DOI 10.7238/rusc.v12i1.1944, no se ha podido establecer algún problema específico por el cual suceda esta falta de obtención de datos.

Tabla 11. Prueba número 2 ejecutada en Linux distribución Debian7

DOI Analizado	Nro de Referencias		Nro de correos obtenidos	
	Total	Con DOI	Automáticamente	Manualmente
10.7238/rusc.v12i1.2031	42	5	3	3
10.7238/rusc.v12i1.1947	29	0	0	0
10.7238/rusc.v12i1.2078	46	12	12	12
10.7238/rusc.v12i1.1944	37	6	0	6
10.7238/rusc.v12i1.1887	31	11	10	10
10.7238/rusc.v12i1.2475	32	2	1	1
10.7238/rusc.v12i1.2253	26	3	3	3
10.7238/rusc.v12i1.2185	31	5	4	4
10.7238/rusc.v12i1.2260	44	5	5	5
10.7238/rusc.v12i1.2262	34	0	0	0
10.7238/rusc.v12i1.2258	19	3	3	3
10.7238/rusc.v12i1.2296	39	3	2	2
10.7238/rusc.v12i1.2446	0	0	0	0

Fuente: Autor.

En la Tabla 12 y Tabla 13 se ejecutó el programa en el sistema operativo Linux distribución Debian 7, obteniendo un tiempo de ejecución para una lista con 13 DOI de 11 minutos 59 segundos para la primera ejecución y 5 minutos 58 segundos para la segunda ejecución.

Tabla 12. Prueba número 1 ejecutada en Windows7

DOI Analizado	Nro de Referencias		Nro de correos obtenidos	
	Total	Con DOI	Automáticamente	Manualmente
10.7238/rusc.v12i1.2031	42	5	3	3
10.7238/rusc.v12i1.1947	29	0	0	0
10.7238/rusc.v12i1.2078	46	12	12	12
10.7238/rusc.v12i1.1944	37	6	6	6
10.7238/rusc.v12i1.1887	31	11	10	10
10.7238/rusc.v12i1.2475	32	2	1	1
10.7238/rusc.v12i1.2253	26	3	3	3
10.7238/rusc.v12i1.2185	31	5	4	4
10.7238/rusc.v12i1.2260	44	5	5	5
10.7238/rusc.v12i1.2262	34	0	0	0
10.7238/rusc.v12i1.2258	19	3	3	3
10.7238/rusc.v12i1.2296	39	3	2	2
10.7238/rusc.v12i1.2446	0	0	0	0

Fuente: Autor.

Tabla 13. Prueba número 2 ejecutada en Windows7

DOI Analizado	Nro de Referencias		Nro de correos obtenidos	
	Total	Con DOI	Automáticamente	Manualmente
10.7238/rusc.v12i1.2031	42	5	3	3
10.7238/rusc.v12i1.1947	29	0	0	0
10.7238/rusc.v12i1.2078	46	12	12	12
10.7238/rusc.v12i1.1944	37	6	6	6
10.7238/rusc.v12i1.1887	31	11	10	10
10.7238/rusc.v12i1.2475	32	2	1	1
10.7238/rusc.v12i1.2253	26	3	3	3
10.7238/rusc.v12i1.2185	31	5	4	4
10.7238/rusc.v12i1.2260	44	5	5	5
10.7238/rusc.v12i1.2262	34	0	0	0
10.7238/rusc.v12i1.2258	19	3	3	3
10.7238/rusc.v12i1.2296	39	3	2	2
10.7238/rusc.v12i1.2446	0	0	0	0

Fuente: Autor.

CONCLUSIONES

- El WebScraping ayuda a recopilar información de sitios web; al aprovechar esta técnica de extracción obtenemos información rápidamente de SCOPUS lo que permite ofrecerla a los lectores a través de métricas al instante.
- Mediante la construcción de un plugin para la visualización automática de métricas en los artículos de OJS, los autores/lectores pueden realizar rápidas inferencias del impacto que está teniendo un artículo. Esta implementación de métricas permite realizar un seguimiento del crecimiento en la web de un artículo desde el momento que es liberado al público.
- Al reorganizar la información de un artículo en OJS, se da al lector facilidad para encontrar cierta información referente al artículo, como por ejemplo ubicar los comentarios o las métricas del artículo en cuestión.
- Mediante e-mail marketing se puede difundir rápidamente y de forma específica un artículo a un grupo seleccionado, lo que permite amplificar el impacto de un artículo dentro de una comunidad académica o científica.

RECOMENDACIONES

- Al momento de realizar un diseño de presentación de publicaciones se debe tener en cuenta la naturaleza de los dispositivos en los que se vaya a presentar la información, un diseño responsivo permite visualizar y transformar las interfaces de presentación de los datos.
- Para obtener la información necesaria de los artículos mediante WebScraping es importante analizar la estructura HTML, identificar las etiquetas y sus atributos y posteriormente establecer un patrón para extraer los datos requeridos.
- Existen varias herramientas para la difusión de artículos científicos y dependiendo del contexto y del público al cual se encuentre dirigido se deber seleccionar aquellas que más se ajuste a la implementación requerida, dentro de este proyecto se ha optado por el e-mail marketing puesto que los correos de destino se pueden extraer fácilmente de la información presentada.
- Se debe buscar alternativas como la propuesta en la herramienta de marketing para evitar caer en malas prácticas de difusión de la información como el envío masivo de SPAM.

Bibliografía

- Canós, J., Letelier, P., & Penadés, M. (12 de Noviembre de 2003). *Ingeniería del Software y Sistemas de Información*. Recuperado el 10 de Enero de 2016, de <http://issi.dsic.upv.es/archives/f-1069167248521/actas.pdf>
- Adie, E., & Roe, W. (Enero de 2013). Altmetric: enriching scholarly content with article-level discussion and metrics. *Learned Publishing*, 26(1), 11-17.
- Alarcia, D. T. (2016). *Nuevo Mundo*. Recuperado el 10 de Enero de 2016, de <http://nuevomundo.revues.org/59836#tocto1n1>
- Bluebill Advisors, Inc. (Marzo de 2003). The classification & evaluation of content management systems. *The Gilbane Report*, 11(2), 3.
- Boiko, B. (2001). Understanding Content Management. *Bulletin of the American Society for Information Science and Technology*, 28, 8-13.
- D'APUZZO, R., & WORWA, B. (2012). Métodos y técnicas para el cálculo y visualización de métricas de historiales en artículos de wikis.
- Davenport, & Prusak. (2001). *Conocimiento en acción. Cómo las organizaciones manejan lo que saben*. Buenos Aires: Prentice Hall.
- DOI®. (Junio de 2014). *Frequently Asked Questions about the DOI® System*. Recuperado el 26 de Julio de 2014, de <http://www.doi.org/faq.html>
- ELIMINAR. (2010). *Open Journal Systems: Una guía completa para la edición de publicaciones en línea*. Recuperado el 2014 de Mayo de 2014, de <http://pkp.sfu.ca/ojs/docs/userguide/2.3.3/es/>
- Eve, M. P. (2011). *DR Martin Paul Eve*. Recuperado el 15 de Mayo de 2014, de <https://www.martineve.com/2011/09/17/writing-a-plugin-for-open-journal-systems-part-1/>
- Fenner, M. (Octubre de 2013). *PLOS Biology*. Recuperado el 10 de Enero de 2016, de <http://www.plosbiology.org/article/fetchObject.action?uri=info:doi/10.1371/journal.pbio.1001687&representation=PDF>
- García, I. M. (2015). *Universidad Pablo de Olavide*. Recuperado el 21 de Enero de 2016, de https://www.upo.es/moleqla/export/sites/moleqla/documentos/Numero19/Nxmero_19.pdf
- Government HKSAR, H. K. (Febrero de 2008). *INFOSEC*. Recuperado el 10 de Enero de 2016, de <http://www.infosec.gov.hk/english/technical/files/web.pdf>
- Hernández, C. (2014). Aplicación de técnicas de web scraping al Boletín Oficial de CAstilla y León (BOCyL).
- Juslén, J. (2011). *Nettimarkkinoinnin karttakirja*. Tietosykli Oy.
- Ling, X., Liu, Y., Huang, Z., Shah, P., & Li, C. (03 de Noviembre de 2015). *Springer Link*. Recuperado el 25 de Enero de 2016, de <http://link.springer.com/article/10.1007%2Fs11192-015-1782-4>
- Martín Galán, B., Rodríguez Mateos, D., Arellano Pardo, M. d., Hernández Pérez, A., & Nogales Flores, J. T. (Abril de 2005). *Gestión de Contenidos Web mediante herramientas de software libre*. Madrid: Unicersidad Carlos III de Madrid.
- Miniwatts Marketing Group. (2015). *Internet World Stats*. Recuperado el Diciembre de 2015, de [internetworldstats.com: http://www.internetworldstats.com/stats.htm](http://www.internetworldstats.com/stats.htm)
- Nakano, R. (2002). *Web Content Management: A Collaborative Approach*. Addison-Wesley.
- Open Journal Systems. (Junio de 2008). *OJS Technical Reference*. Recuperado el 30 de Enero de 2015, de Open Journal Systems: <http://pkp.sfu.ca/ojs/OJSTechnicalReference.pdf>
- Oulu, U. o. (27 de Noviembre de 2012). *JULTIKA University of Oulu repository*. Recuperado el 10 de Enero de 2016, de <http://jultika.oulu.fi/files/nbnfioulu-201302121033.pdf>
- Penman, R., Baldwin, T., & Martinez, D. (2009). *Web scraping made simple with sitedscraper*.

- Porrúa García, M. (Enero de 2010). *Dintel*. Recuperado el 26 de Enero de 2016, de <http://www.revistadintel.es/Revista/Numeros/Numero4/TyG/porrua.pdf>
- Pride, W., & Ferrell, O. (2012). *Ramsdell Design*. Recuperado el 21 de Enero de 2016, de http://www.ramsdelldesign.com/pdf/pride_ccd.pdf
- Public Knowledge Project. (2010). Recuperado el 10 de Enero de 2016, de <https://pkp.sfu.ca/ojs/OJSTechnicalReference.pdf>
- Robertson, J. (2003). *So, what is a content management system?* Recuperado el 29 de Mayo de 2014, de http://www.steptwo.com.au/files/kmc_what.pdf
- Rojas, M. Y. (16 de Febrero de 2006). *BVSCUBA*. Recuperado el 21 de Enero de 2016, de http://bvs.sld.cu/revistas/aci/vol14_1_06/aci02106.htm
- Serradell, E., & Juan, Á. (Febrero de 2003). *Universitat Oberta de Catalunya*. Recuperado el 21 de Enero de 2016, de <http://www.uoc.edu/dt/20133/20133.pdf>
- Suber, P. (2012). *mitpress*. Recuperado el 10 de Enero de 2016, de https://mitpress.mit.edu/sites/default/files/9780262517638_Open_Access_PDF_Version.pdf
- Sweeney, S., MacLellan, A., & Dorey, E. (2006). *3G Marketing on the Internet: Third Generation Internet Marketing Strategies for Online Success*. Maximum Press.
- Turban, E., King, D., Lee, J., Liang, T.-P., & Turban, D. (2010). *Gemeinsamer Bibliotheksverbund*. Recuperado el 21 de Enero de 2016, de <http://www.gbv.de/dms/zbw/608075914.pdf>
- Turban, E., Lee, J., King, D., Liang, T., & Turban, D. (2010). *Electronic Commerce 2010: A Managerial Perspective*. New Jersey: Prentice Hall.

ANEXOS

A. Configuraciones iniciales.

Para la realización de este proyecto se necesitará configurar un ambiente de desarrollo para OJS, estas configuraciones se realizarán sobre el sistema operativo Linux en una de sus distribuciones más estables llamada Debian versión 7 que tiene como nombre "Wheezy" en la cual se configurará los siguientes programas:

a. PHP

Debian tiene repositorios de paquetes o dichos de otra manera instaladores de programas accesibles desde Internet; gracias a estos repositorios la instalación de la mayoría de programas es sumamente sencilla.

Se puede acceder a los programas en los repositorios de 2 maneras diferentes pero sencillas.

La primera manera de instalación es a través del gestor de paquetes llamado Synaptic, en este caso no es importante conocer exactamente que paquete necesitamos, pero se debe saber que programa necesitamos que sea instalado y basta con buscar dicho nombre ya sea a través del filtro o manualmente y marcarlo para ser instalado; luego de marcar para instalar, automáticamente nos sugerirá instalar otros paquetes de dependencia los cuales son necesarios para que el programa (en este caso PHP5) funcione correctamente.

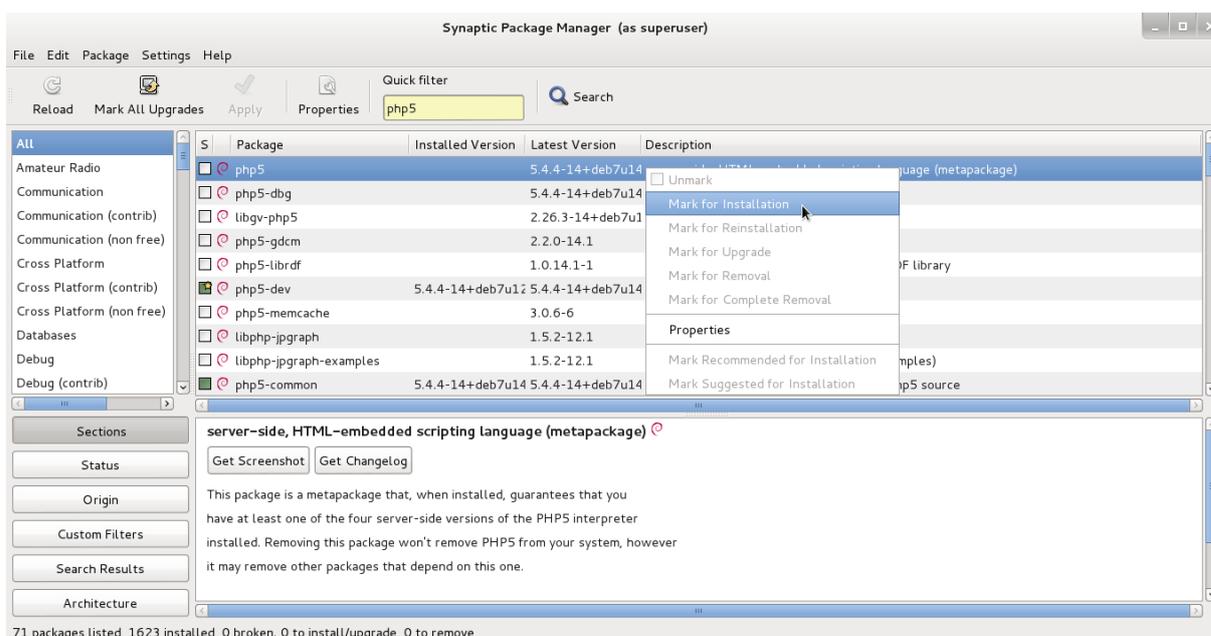
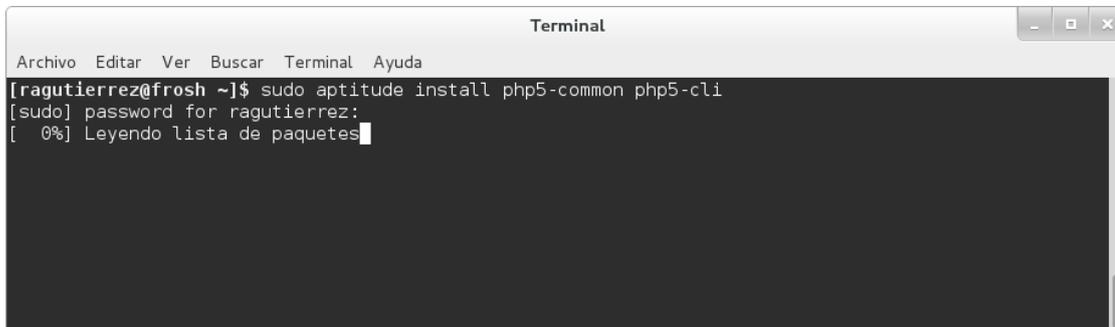


Figura 37: Instalación de PHP en Synaptic.

Fuente: Autor.

La segunda opción es un poco más avanzada ya que requiere el uso del terminal (consola de comandos de Linux) y también debemos saber los nombres exactos de los paquetes que deseamos instalar.



```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
[ragutierrez@frosh ~]$ sudo aptitude install php5-common php5-cli
[sudo] password for ragutierrez:
[ 0%] Leyendo lista de paquetes
```

Figura 38: Comando de instalación para PHP5 en la consola (terminal).
Fuente: Autor.

b. MySQL

Al igual que PHP, MySQL se puede instalar de dos maneras diferentes gracias a los repositorios de Debian que nos hacen la vida mucho más fácil.

A través de Synaptic:

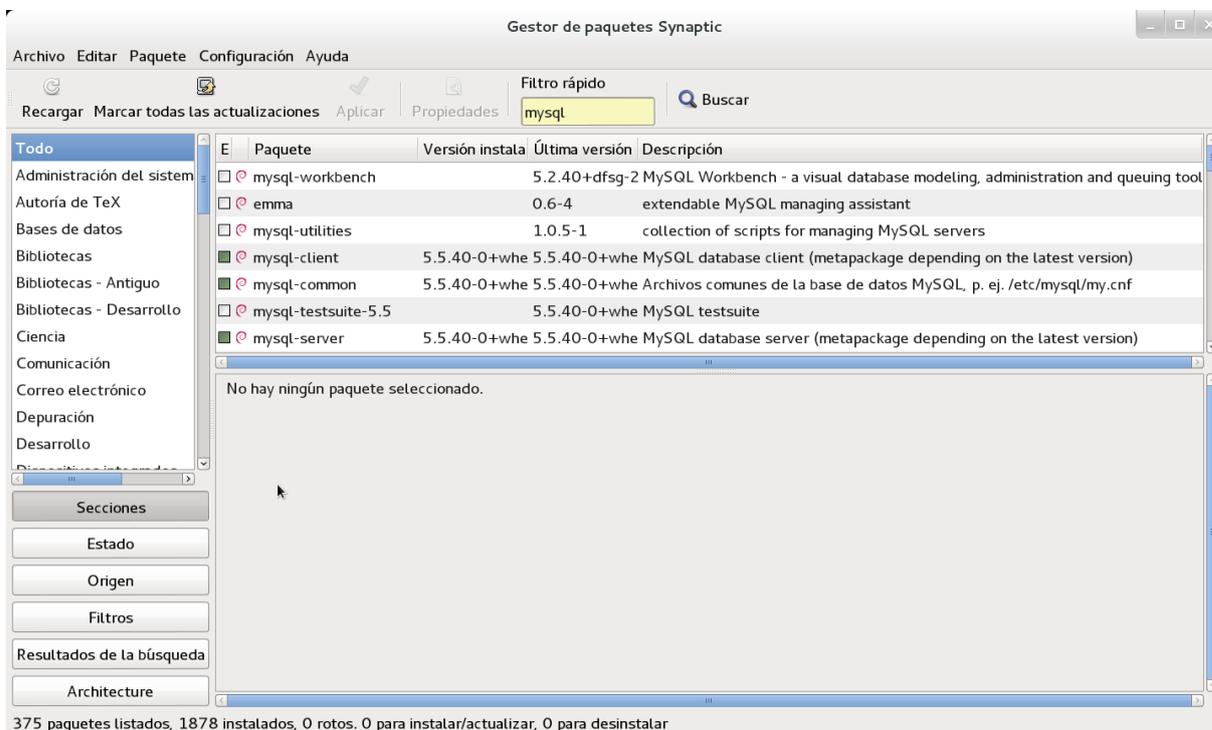
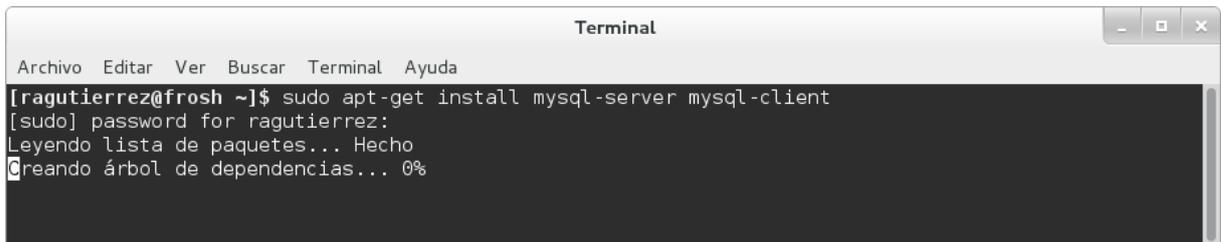


Figura 39: Instalación de MySQL a través de Synaptic.
Fuente: Autor.

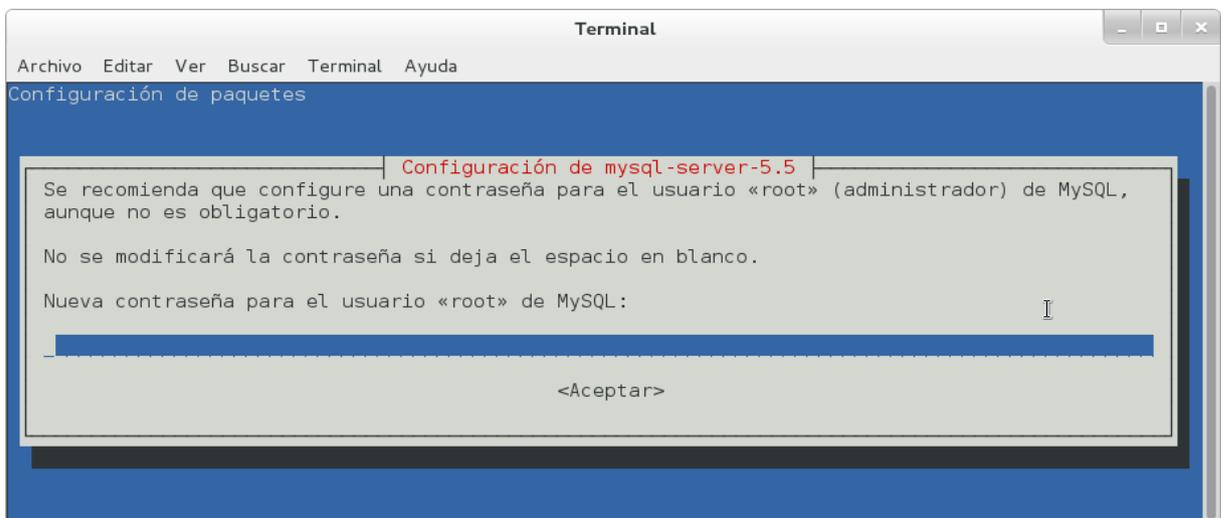
O también a través de comandos en el terminal:



```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
[ragutierrez@frosh ~]$ sudo apt-get install mysql-server mysql-client
[sudo] password for ragutierrez:
Leyendo lista de paquetes... Hecho
Creando árbol de dependencias... 0%
```

Figura 40: Comandos de instalación de MySQL a través del Terminal.
Fuente: Autor.

Independientemente de la forma de instalación, MySQL lanzará una configuración inicial en la cual podremos introducir la contraseña de seguridad para nuestro servidor de bases de datos.



```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
Configuración de paquetes
Configuración de mysql-server-5.5
Se recomienda que configure una contraseña para el usuario «root» (administrador) de MySQL, aunque no es obligatorio.
No se modificará la contraseña si deja el espacio en blanco.
Nueva contraseña para el usuario «root» de MySQL:
[Redacted password input]
<Aceptar>
```

Figura 41: Petición de contraseña para MySQL Server 5.5.
Fuente: Autor.

Como método de seguridad MySQL nos solicita una confirmación de la contraseña, esto es muy útil para evitar errores al ingresar la información.

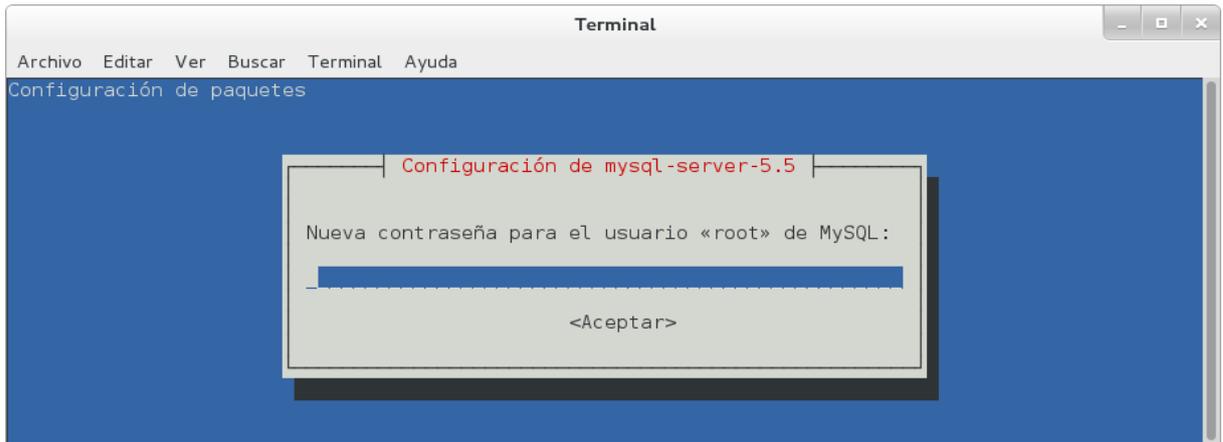


Figura 42: Confirmación de la contraseña introducida para MySQL Server 5.5
Fuente: Autor.

Luego de instalar el sistema de gestión de base de datos, debemos crear una base de datos para ser utilizada por Open Journal Systems. Para ello debemos ingresar a MySQL, el usuario creado por defecto es “root” y la contraseña es la que habíamos creado anteriormente; una vez dentro del programa, creamos la base de datos mediante la sintaxis SQL. Si todo ha ido bien, deberíamos tener un mensaje de respuesta que diga “Query OK, 1 row affected”.

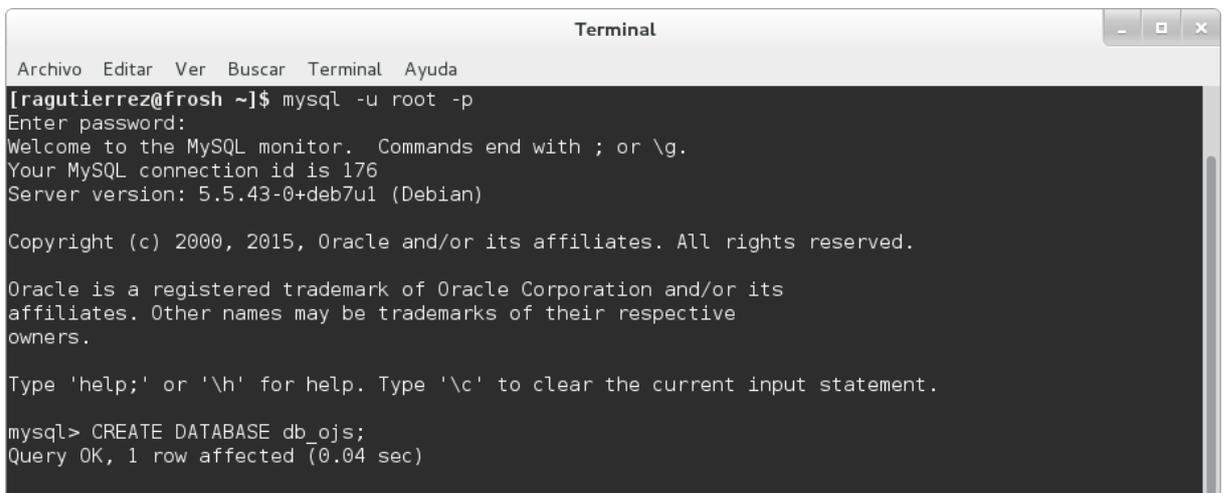
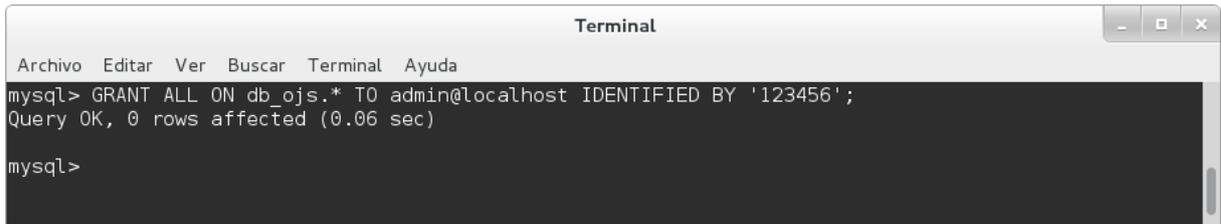


Figura 43: Creación de la base de datos para OJS.
Fuente: Autor.

OJS recomienda la creación de un nuevo usuario exclusivo para la base de datos a utilizar al cual se le otorgarán todos los permisos para todas las tablas que se creen dentro de la base de datos; y lo realizamos utilizando la sintaxis SQL. La sentencia será la siguiente:



```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
mysql> GRANT ALL ON db_ojs.* TO admin@localhost IDENTIFIED BY '123456';
Query OK, 0 rows affected (0.06 sec)

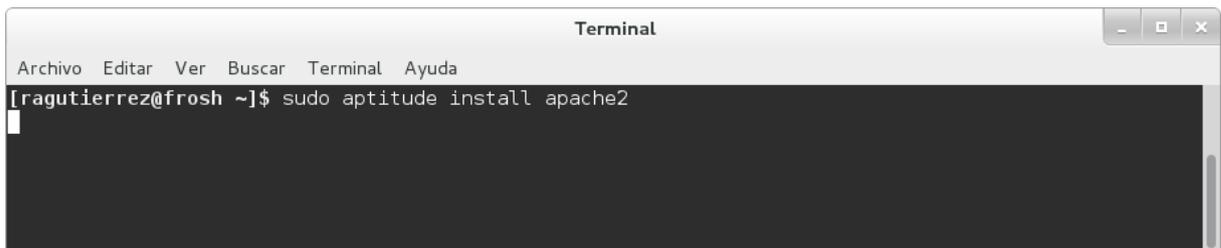
mysql>
```

Figura 44: Asignación de permisos a un usuario para la base de datos de OJS.
Fuente: Autor.

c. Apache

La instalación de apache se la puede hacer también a través de los repositorios de Debian o mediante la descarga de los archivos fuente. Para esta ocasión hemos de instalar a través de los repositorios de Debian con el fin de que no exista una mala configuración manual.

Esta vez solo utilizaremos la consola de comandos o terminal para la instalación en la cual ingresaremos los siguientes comandos:



```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
[ragutierrez@frosh ~]$ sudo aptitude install apache2
```

Figura 45: Comando de instalación de apache.
Fuente: Autor.

Para comprobar que la instalación de Apache fue correcta, ingresamos al navegador la URL local o en su defecto <http://localhost/> y debería mostrar el siguiente mensaje:

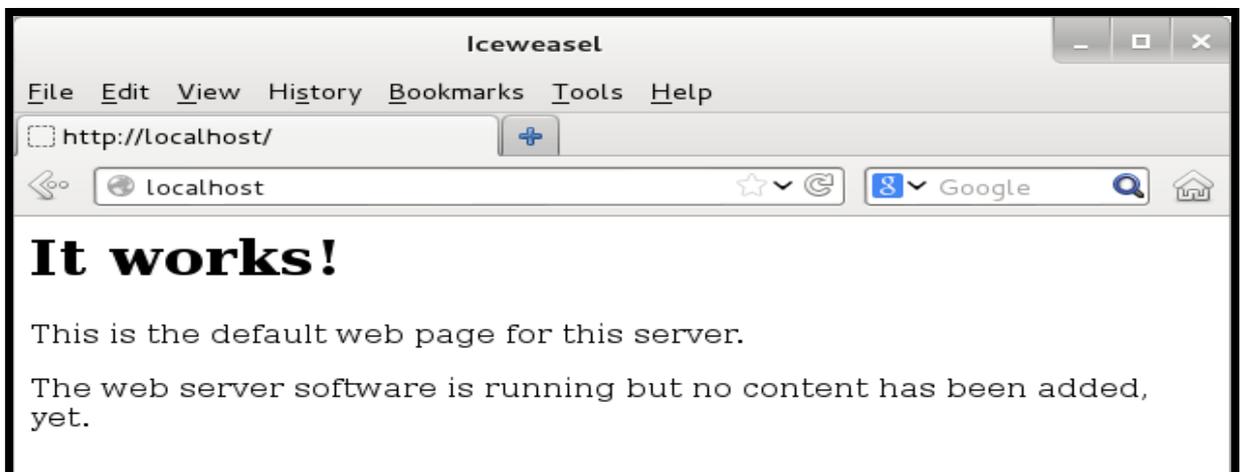


Figura 46: Captura del mensaje que demuestra que apache2 está instalado.
Fuente: Autor.

Una vez que comprobamos que Apache está correctamente instalado, debemos comprobar si PHP se puede integrar sin ninguna dificultad. Para ello creamos en la carpeta web de apache un archivo llamado “phpinfo.php” o con cualquier otro nombre, el único requisito es que contenga la extensión “.php” y que este archivo contenga en su interior lo siguiente:

```
1 <?php
2     phpinfo();
3 ?>
4
```

Figura 47: Archivo de verificación de php
Fuente: Autor.

Al ingresar desde el navegador al archivo creado nos debería aparecer lo siguiente:

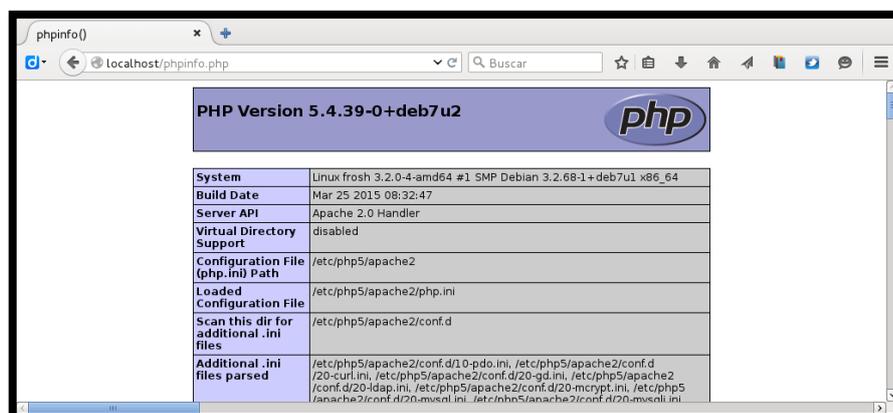


Figura 48: información de PHP al ingresar en el navegador al archivo de información.
Fuente: Autor.

d. OJS

La instalación de Open Journal Systems es relativamente fácil, se necesita de unas configuraciones sencillas que el mismo OJS las anuncia mientras se lo instala. Antes que nada debemos descargar el paquete estable que contiene los archivos de OJS a través de su página:

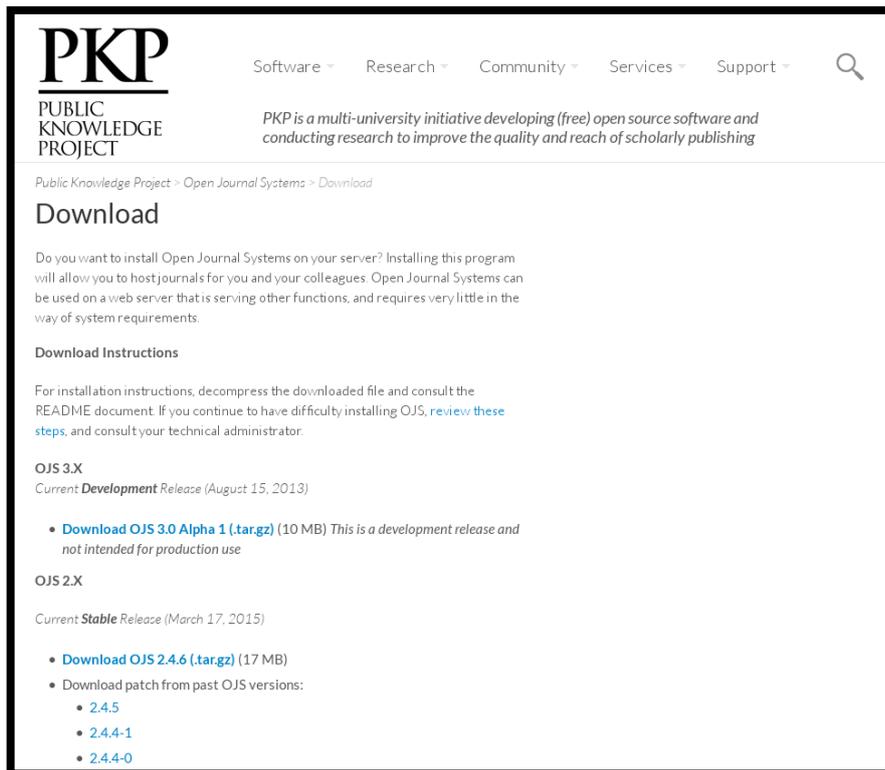


Figura 49: Página de PKP para la descarga de OJS.

Fuente: http://pkp.sfu.ca/ojs/ojs_download/

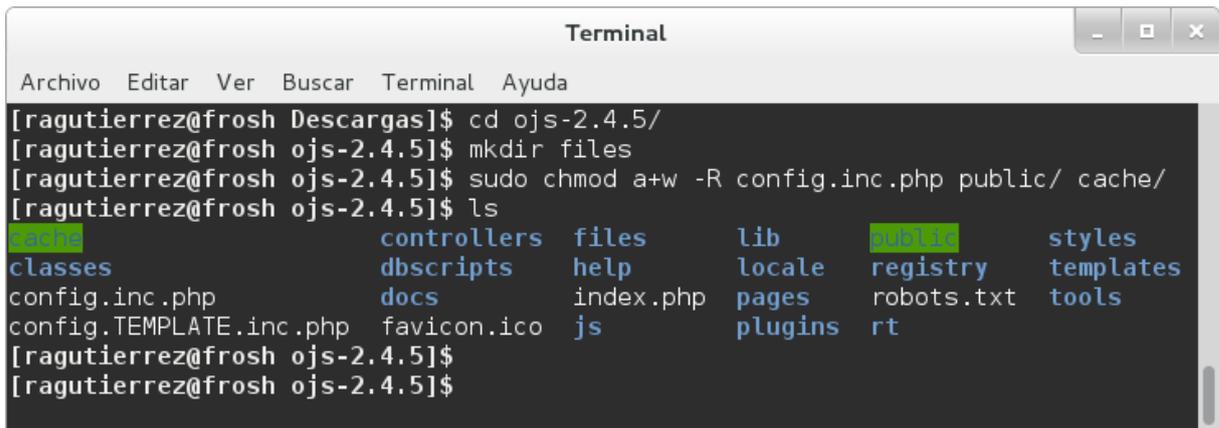
El paquete descargado deberá ser descomprimido para poder ser utilizado. En el sistema operativo Linux es relativamente fácil ya que posee el compresor TAR instalado por defecto.



Figura 50: Descompresión del archivo que contiene OJS.

Fuente: Autor.

Una vez que se ha descomprimido el archivo, ingresamos a la carpeta que se ha creado en este caso “ojs-2.4.5”. OJS nos pide una carpeta para los archivos de los artículos, en este caso crearemos una llamada “files”.



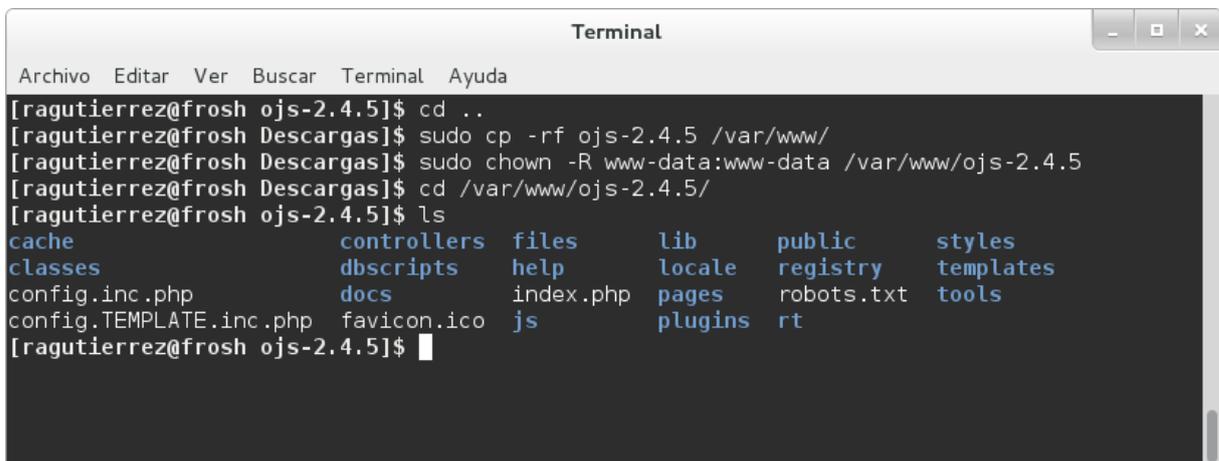
```
Terminal
Archivo  Editar  Ver  Buscar  Terminal  Ayuda
[ragutierrez@frosh Descargas]$ cd ojs-2.4.5/
[ragutierrez@frosh ojs-2.4.5]$ mkdir files
[ragutierrez@frosh ojs-2.4.5]$ sudo chmod a+w -R config.inc.php public/ cache/
[ragutierrez@frosh ojs-2.4.5]$ ls
cache          controllers  files       lib          public       styles
classes       dbscripts   help        locale      registry     templates
config.inc.php docs        index.php   pages       robots.txt   tools
config.TEMPLATE.inc.php favicon.ico js          plugins     rt
[ragutierrez@frosh ojs-2.4.5]$
[ragutierrez@frosh ojs-2.4.5]$
```

Figura 51: comandos para creación de carpetas y permisos de archivos.

Fuente: Autor.

Otra configuración previa antes de comenzar la instalación del sistema es el cambio de permisos a las carpetas “public” y “cache” además del archivo “config.inc.php”, se les otorgará permisos de escritura ya que son necesarios para la instalación.

Una vez configurado, copiaremos la carpeta de OJS hacia la carpeta de sitios web de apache y cambiaremos el usuario “dueño” de la carpeta por el usuario “www-data” ya que es el usuario que normalmente se ejecuta bajo apache.



```
Terminal
Archivo  Editar  Ver  Buscar  Terminal  Ayuda
[ragutierrez@frosh ojs-2.4.5]$ cd ..
[ragutierrez@frosh Descargas]$ sudo cp -rf ojs-2.4.5 /var/www/
[ragutierrez@frosh Descargas]$ sudo chown -R www-data:www-data /var/www/ojs-2.4.5
[ragutierrez@frosh Descargas]$ cd /var/www/ojs-2.4.5/
[ragutierrez@frosh ojs-2.4.5]$ ls
cache          controllers  files       lib          public       styles
classes       dbscripts   help        locale      registry     templates
config.inc.php docs        index.php   pages       robots.txt   tools
config.TEMPLATE.inc.php favicon.ico js          plugins     rt
[ragutierrez@frosh ojs-2.4.5]$ █
```

Figura 52: Copia de archivos PJS a la carpeta de Apache y cambio de usuario de archivos.

Fuente: Autor.

Todo el proceso de instalación se lo realiza a través del navegador. Y para ello debemos ingresar a la URL local añadiendo el nombre de la carpeta así “http://localhost/ojs-2.4.5” el cual nos debería enviar a la siguiente página para la instalación:

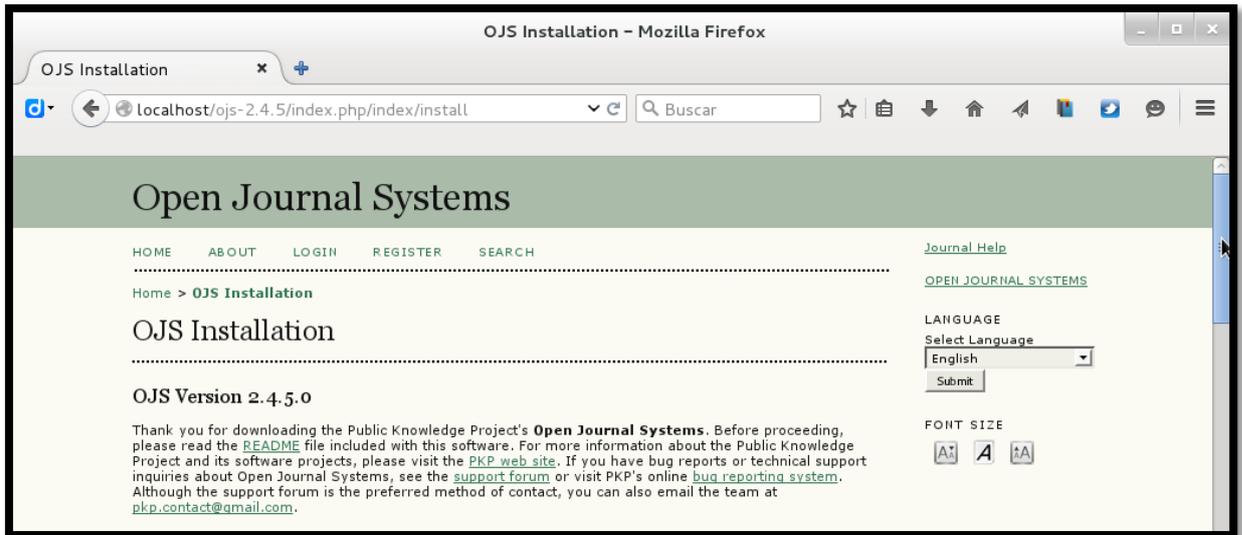


Figura 53: Página de instalación de OJS.
Fuente: Autor.

Si todo ha salido bien, deberíamos comenzar a instalar sin ningún problema rellorando el formulario que se muestra en dicha página. Si existe alguna duda de cómo proceder, siempre se puede acudir a la guía de instalación publicada por el propio PKP en la siguiente URL: <http://pkp.sfu.ca/ojs/docs/userguide/2.3.1/systemAdministrationInstallProcess.html>

B. Ejemplo de archivo CSV que se puede exportar con la herramienta de marketing.

DOI_Principal	Título_Artículo_Principal	Autor_Artículo_Principa	DOI_Secundario	Título_Artículo_Secundario	Autor_Artículo_Sec	E-mail
10.7238/rusc.v12i1.2031	Model for democratisation of	Atenas, J.	10.1080/00048623.2013.821048	MOOCs: The Challenges for Academic Libraria	Barnes, C.	cbarnes@une.edu.au
10.7238/rusc.v12i1.2031	Model for democratisation of	Atenas, J.	10.7238/rusc.v10i1.1689	New informal ways of learning: Or are we for	Sangrã, A.	asangra@uoc.edu
10.7238/rusc.v12i1.2031	Model for democratisation of	Atenas, J.	10.3916/C38-2012-02-06	«Prosuming» across cultures: Youth creating a	Dezuanni, M.	m.dezuanni@qut.edu.au
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1023/A:1011087813505	Traditional, interpretive, and reception based	Ahuvia, A.	Ahuvia@umich.edu
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1111/j.1365-2729.2010.00369.x	Debunking the 'digital native': Beyond digital	Brown, C.	cheryl.brown@uct.ac.za
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1016/j.sbspro.2011.11.174	Social network sites usage habits of undergra	Bicen, H.	huseyin_bicen@hotmail.com
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1191/1478088706qp063oa	Using thematic analysis in psychology	Braun, V.	v.braun@auckland.ac.nz
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1111/j.1467-8535.2007.00793.x	The 'digital natives' debate: A critical review	Bennett, S.	sue_bennett@uow.edu.au
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1016/j.chb.2010.03.024	Facebook and academic performance	Kirschner, P.A.	paul.kirschner@ou.nl
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1080/17439884.2010.531024	Technological diversity: An investigation of st	Corrin, L.	lcorrin@uow.edu.au
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1080/01411920902989227	Digital natives: Where is the evidence?	Helsper, E.J.	Ellen.helsper@oii.ox.ac.uk
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1111/j.1365-2729.2007.00268.x	Informal learning with PDAs and smartphone	Clough, G.	g.m.clough@open.ac.uk
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1111/j.1365-2729.2010.00371.x	Beyond natives and immigrants: Exploring typ	Kennedy, G.	gek@unimelb.edu.au
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1111/nhs.12048	Content analysis and thematic analysis: Impli	Turunen, H.	hannele.turunen@uef.fi
10.7238/rusc.v12i1.2078	Students in higher education:	Gallardo Echenique, E.	10.1111/j.1467-8535.2011.01195.x	Using the Facebook group as a learning manag	Wang, Q.	qiyun.wang@nie.edu.sg
10.7238/rusc.v12i1.1944	Mobile learning: Perspectives	Torres Diaz, J.C.	10.1109/38.963459	Recent advances in augmented reality	Azuma, R.	azuma@hrl.com
10.7238/rusc.v12i1.1944	Mobile learning: Perspectives	Torres Diaz, J.C.	10.1016/j.sbspro.2011.12.171	Pedagogical framework of m-learning	Ozdamli, F.	fezileozdamli@hotmail.com
10.7238/rusc.v12i1.1944	Mobile learning: Perspectives	Torres Diaz, J.C.	10.3402/rlt.v20i0/14406	Viewing mobile learning from a pedagogical	Kearney, M.	matthew.kearney@uts.edu.au

Figura 54: Ejemplo de archivo CSV obtenido con la herramienta de marketing.

Fuente: Autor.

C. Manual de instalación de la herramienta de marketing.

Se ha llamado al programa “ArtInfo”, se ha creado este proyecto para obtener los correos de las referencias de un artículo mediante el servicio que ofrece SCOPUS.

Una vez recolectados todos los correos, se dará una opción al usuario para que los exporte a un archivo CVS.

a. Requisitos: En el ordenador a instalar, se deben tener los siguientes programas.

- Python (2.7)
- Qt (4.8)
- MySQL (5)

b. Configuraciones:

- Configuración para Windows
 - Agregar la siguiente línea a la variable de entorno 'Path':
"C:\Python27\;C:\Python27\Scripts\"
 - Instalar Compilador C++ para python 2.7
El instalador lo puede encontrar en la siguiente dirección web:
<http://www.microsoft.com/en-us/download/confirmation.aspx?id=44266>
 - Instalar MySQL Connector para python
El instalador lo puede encontrar en la siguiente dirección web:
<http://dev.mysql.com/downloads/connector/c/6.0.html#downloads>
- Para todos los sistemas:
 - Instalar los paquetes Python “pip” y “virtualenv”.
- Instalación
- Verificar que los programas detallados en “Requisitos” se encuentren previamente instalados.
- Descargar los archivos del siguiente repositorio web:
<https://git.taw.utpl.edu.ec/ragutierrez/ArtInfo/repository/archive.zip>
- Descomprimir el archivo descargado.
- Crear un entorno virtual con el nombre “qt” en la carpeta “env”
- Activamos el entorno virtual.
- Instalar con el comando “pip” los requerimientos en el archivo “requirements.txt” dentro de la carpeta “config”.
- Crear una base de datos en “mysql” con los siguientes comandos:
 - CREATE DATABASE IF NOT EXISTS db_artinfo DEFAULT CHARACTER SET utf8;

```
- GRANT ALL ON db_artinfo.* TO
  usuario_secundario@localhost IDENTIFIED BY
  'contraseña_del_usuario';
```

- Es importante cambiar "usuario_secundario" y "contraseña_del_usuario" por datos adecuados.
- Obtener una APIKEY de SCOPUS en la siguiente página web:
<http://dev.elsevier.com>
- Editar el archivo "config.ini" en la carpeta "config".
- Es importante cambiar únicamente las líneas 4, 5 y 9 que corresponden al usuario, contraseña y APIKEY; por los creados en los pasos anteriores.

c. Recomendaciones.

- Desactivar/Deshabilitar IPv6 (suele causar problemas con el funcionamiento de PhantomJS)

D. Plugin. Manual de Programador

a. Estructura de archivos de OJS

Es importante conocer la estructura de archivos que OJS provee con el fin de tener un conocimiento de en qué archivo vamos a intervenir.

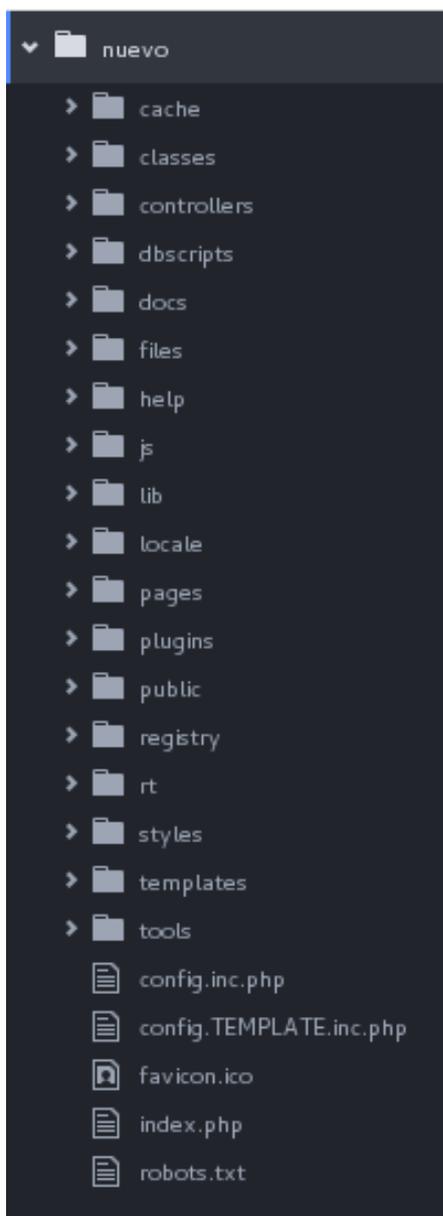


Figura 55: Carpetas y archivos de OJS

Fuente: Autor.

OJS tiene una serie de archivos y subcarpetas distribuidos en carpetas generales (Figura 55) que para su mejor entendimiento se agrupan por el desempeño de los archivos; se detallarán algunas para no extender mucho este apartado. Las carpetas “classes” y “lib” contienen los archivos escritos en el lenguaje de programación web “php” para que nuestra

instalación OJS funcione correctamente y se pueda armar el sitio para administrar nuestras revistas en línea y los artículos dentro de las mismas.

Otra de las carpetas de alta importancia en OJS es la carpeta “plugins”, dentro de esta se encuentran instalados por defecto algunas de las extensiones que el equipo de desarrolladores de PKP ha considerado serán de mucha utilidad y que se podrían utilizar en las revistas; esta carpeta también está destinada a almacenar las extensiones que desarrolladores independientes puedan crear para mejorar sus propias instalaciones de OJS. Finalmente, en la carpeta “templates” se han colocado las plantillas para que OJS cargue la información de los artículos.

b. Procedimiento para crear el Plugin

En OJS los plugin sirven para cambiar y/o extender el comportamiento de las plantillas sin necesidad de cambiar el código base así las demás revistas no se verán afectadas por dichos plugin. Los plugin están pensados para realizar extensiones o modificaciones a OJS a través de una colección de código y recursos que deben seguir una estructura definida por el mismo OJS.

Para la implementación del proyecto se tomó como guía las especificaciones presentadas en (Eve, 2011) el cual es un tutorial que detalla lo siguiente:

1) Crear una carpeta con archivos básicos.

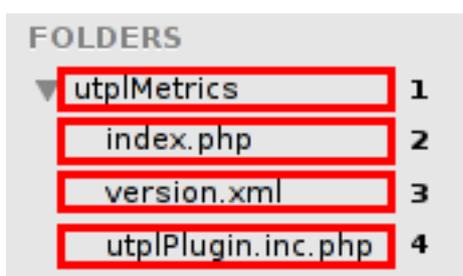


Figura 56: Archivos básicos dentro de la carpeta "utplMetrics".

Fuente: Autor.

Se sugiere que el nombre de la carpeta tenga relación con la finalidad del plugin ya que esta servirá para empaquetar todos los archivos que consideremos necesario crear para que el plugin funcione adecuadamente. Una vez finalizado el desarrollo, esta carpeta deberá estar comprimida en el formato “.tar.gz” para que pueda ser instalada mediante el gestor de plugins en OJS.

El archivo `index.php`, aunque vital, sirve únicamente para hacer referencia hacia nuestro archivo principal del plugin donde se encuentra la codificación principal; dicho archivo solo contiene lo que se muestra en la Figura 57.

```
index.php x
1  <?php
2
3  /**
4   * @defgroup plugins_generic_utplPlugin
5   */
6
7  /**
8   * @file plugins/generic/utplMetrics/index.php
9   *
10  */
11
12
13  require_once('utplPlugin.inc.php');
14
15  return new utplPlugin();
16
17  ?>
```

Figura 57: archivo “index.php”.

Fuente: Autor.

Otro archivo básico para la creación es el “versión.xml”, este archivo contiene la información básica del control de versiones de nuestro plugin y OJS lo tomará como guía para instalarlo.

```
version.xml x
1  <?xml version="1.0" encoding="UTF-8"?>
2  <!DOCTYPE version SYSTEM "../..../..
3  /lib/pkp/dtd/pluginVersion.dtd">
4
5  <version>
6     <application>utplMetrics</application>
7     <type>plugins.generic</type>
8     <release>1.0.0.0</release>
9     <date>2014-11-27</date>
10    <lazy-load>1</lazy-load>
11    <class>utplPlugin</class>
12 </version>
```

Figura 58: archivo “version.xml”

Fuente: Autor.

El cuarto es el archivo más importante de la lista puesto que en él se encuentra toda la codificación necesaria para cumplir un objetivo con el plugin, en este caso el de enriquecer el contenido de los artículos mediante métricas. En la Figura 59 solo se muestran las primeras líneas del extenso código.

```
utplPlugin.inc.php x
1  <?php
2  /*
3  *
4  * archivo: utplPlugin.inc.php
5  * se ha tomado como referencia: ALM Plugin y Google
6  *
7  */
8
9  import('lib.pkp.classes.plugins.GenericPlugin');
10 import('lib.pkp.classes.core.JSONManager');
11
12 DEFINE('ALTMETRIC_URL', 'http://api.altmetric.com/v1');
13 DEFINE('ELSEVIER_URL', 'http://api.elsevier.com/content');
14
15 class utplPlugin extends GenericPlugin {
16
17     /** @var $altmetricKey string */
18     var $_altmetricKey="";
19     /** @var $scopusKey string */
20     var $_scopusKey="";
21
22     /**
23      * @see LazyLoadPlugin::register()
24      */
25     function register($category, $path) {
26         $success = parent::register($category, $path);
27         if (!Config::getVar('general', 'installed')) {
28
29             $application =& Application::getApplication();
30             $request =& $application->getRequest();
31             $router =& $request->getRouter();
32             $context = $router->getContext($request);
```

Figura 59: archivo "utplPlugin.inc.php".

Fuente: Autor.

2) Identificar el Hook al cual se deba añadir el desarrollo.

Los Hooks (ganchos) son usados por los plugin como una herramienta de notificación y para modificar comportamientos preestablecidos en OJS. Como habíamos dicho anteriormente, OJS se maneja a través de plantillas las cuales podremos modificar a nuestra conveniencia para darle otra presentación o visualización.

En la guía de referencias técnicas (Open Journal Systems, 2008) existe una lista de Hooks (con sus nombres, parámetros y descripciones) la cuál es de mucha ayuda para ubicar de manera precisa el elemento que debemos utilizar.

<code>_returnSubscriptionTypeFromRow</code>	<code>&\$row</code>	Called before the template manager (<code>&\$templateMgr</code>) sends the content type header with the given content type (<code>&\$sendContentType</code>) and character set (<code>&\$charset</code>) and displays a template (<code>&\$template</code>). To prevent OJS from performing this action, the hook registrant should return <code>true</code> from its callback.
<code>TemplateManager::display</code>	<code>&\$templateMgr, &\$template, &\$sendContentType, &\$charset</code>	Called after <code>UserDAO</code> builds a <code>User</code> (<code>&\$user</code>) object from the database row
<code>UserDAO::returnUserFromRow</code>	<code>&\$user, &\$row</code>	

Figura 60: Hook "Template Manager::display".

Fuente: plugins Hook List en (Open Journal Systems, 2008)

El hook que sirve para manejar las plantillas se llama "TemplateManager::display", a este hook se lo llamará para gestionar las nuevas plantillas que se crearán a partir de las existentes en OJS.

c. Consumo de ELSEVIER-SCOPUS

El servicio de ELSEVIER es exclusivo para entidades que hayan comprado sus prestaciones o para entidades que para fines de investigación soliciten acceso a dichos servicios.

Una segunda forma de acceder al servicio de búsqueda en su base de datos de citas SCOPUS, es a través de la solicitud de una API Key desde la página de desarrolladores de ELSEVIER (<http://dev.elsevier.com/>), desde la cual se puede acceder sin necesidad de estar dentro de una red corporativa o haber pagado para el uso de sus servicios.

El API Key sirve para autenticar el software que se está conectando con la base de datos SCOPUS y así evitar ataques de denegación de servicios a través del API.

Para obtener el API Key de ELSEVIER, hay que registrarse a través de su web para desarrolladores en la cual hay que llenar un pequeño formulario con información personal similar al que se rellena para crear cuentas en las redes sociales (Figura 61). Una vez creada la cuenta, deberemos confirmarla a través del correo y procedemos a ingresar en la página para generar la API Key.

ELSEVIER Developers Overview Use Cases Explore My API key API Specification

My Projects Register | Login

Register

Registration is quick and free. It allows you to personalize these [Elsevier Products](#) if you have access. For example you can stay up-to-date with Search Alerts and Document Citation Alerts or keep track of your research with Saved Searches.

(* = required fields)

Your details [Privacy policy](#)

First name: *

Family name: *

E-mail and password

Enter a password between 5 and 20 characters. Your e-mail address will be your username.

E-mail address: *

Password: *

Confirm password: *

Your organization, role and field of interest

Organization:

Figura 61: formulario para crear una cuenta de ELSEVIER Developers.

Fuente: <http://dev.elsevier.com/>

Al acceder con la cuenta creada, veremos un recuadro en la parte derecha en el cual mediante un botón, nos permitirá registrar una aplicación para que se pueda conectar al servicio de SCOPUS y realizar las consultas que se consideren pertinentes. Una vez que se llena un pequeño formulario con información acerca de la aplicación, veremos en la parte izquierda los datos ingresados y un enlace que nos mostrará nuestro API Key (Figura 62).

ELSEVIER Developers Overview Use Cases Explore My API key api.elsevier.com

My Projects Rommel Augusto Gutierrez Roa

The My Project page provides you with a list of your applications and registered websites.

My Registered Websites/Applications		Actions
Registered on	Website URL	
15 Oct 2014	http://none	View API Key Edit Delete

Register a new site

Register your website or application here!

[Register a New Site](#)

Figura 62: "My Projects", sección para obtener la API Key.

Fuente: <http://dev.elsevier.com/>

Dentro del servicio de búsqueda de SCOPUS, mediante su herramienta HTML, existen varias opciones para la búsqueda de datos, se puede solicitar información por atributos generales como: autores, editoriales, títulos de artículos, títulos de revistas y muchos más; pero también están la búsqueda a través de identificadores únicos como el ISSN (International Standard Serial Number; Número Internacional Normalizado de Publicaciones Seriadas), el CODEN (código alfanumérico de seis caracteres) o el DOI (Digital Object Identifier).

Mediante el API de ELSEVIER podemos acceder a las mismas opciones de búsqueda pero con respuestas más específicas en formato XML. La búsqueda de citas mediante el API que provee ELSEVIER, posee la debida documentación enfocada a desarrolladores disponible en la siguiente dirección web:

<http://api.elsevier.com/documentation/SCOPUSSearchAPI.wadl>

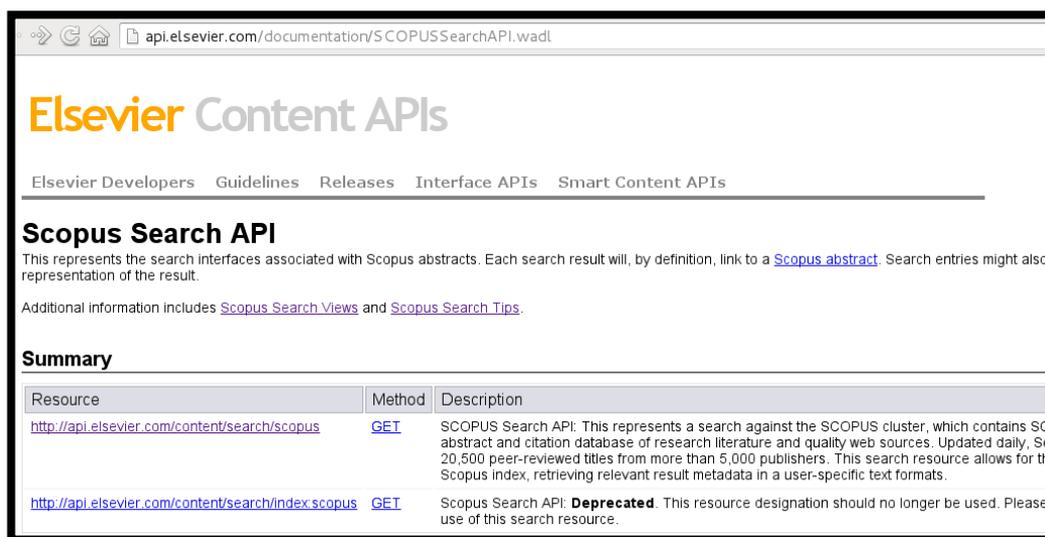


Figura 63: Documentación en línea del API para la búsqueda en SCOPUS.

Fuente: Autor.

Con la ayuda de dicho manual se puede saber que la sintaxis de la URL para la búsqueda de información en SCOPUS es de la siguiente manera:

API URL + API Key + DOI



Figura 64: URL formada para la consulta del DOI en el API de Altmetric.

Fuente: Autor.

Cuando toda esta cadena de caracteres está bien formada, el navegador nos devuelve la información en formato XML, el cual se puede leer para obtener las etiquetas que sean más relevantes. Para efectos de este capítulo, la única etiqueta que nos interesa es la que hace referencia al conteo de citas que sería: “<citedby-count></citedby-count>”.

```
▼<search-results xmlns="http://www.w3.org/2005/Atom" xmlns:atom="http://www.w3
xmlns:opensearch="http://a9.com/-/spec/opensearch/1.1/" xmlns:prism="http://p
<opensearch:totalResults>1</opensearch:totalResults>
<opensearch:startIndex>0</opensearch:startIndex>
<opensearch:itemsPerPage>1</opensearch:itemsPerPage>
<opensearch:Query role="request" searchTerms="DOI%28%2210.1371%2Fjournal.pone
<link href="http://api.elsevier.com:80/content/search/scopus?
start=0&count=25&query=DOI%28%2210.1371%2Fjournal.pone.0069841%22%29apiKey=3
<link href="http://api.elsevier.com:80/content/search/scopus?
start=0&count=25&query=DOI%28%2210.1371%2Fjournal.pone.0069841%22%29apiKey=3
▼<entry>
<link href="http://api.elsevier.com/content/abstract/scopus_id/84881488287
<link href="http://www.scopus.com/inward/record.url?partnerID=Hz0xMe3b&sc
<link href="http://www.scopus.com/inward/citedby.url?partnerID=Hz0xMe3b&sc
▼<prism:url>
http://api.elsevier.com/content/abstract/scopus_id/84881488287
</prism:url>
<dc:identifier>SCOPUS_ID:84881488287</dc:identifier>
<eid>2-s2.0-84881488287</eid>
▼<dc:title>
Facebook Use Predicts Declines in Subjective Well-Being in Young Adults
</dc:title>
<dc:creator>Kross, E.</dc:creator>
<prism:publicationName>PLoS ONE</prism:publicationName>
<prism:eIssn>19326203</prism:eIssn>
<prism:volume>8</prism:volume>
<prism:issueIdentifier>8</prism:issueIdentifier>
<prism:coverDate>2013-08-14</prism:coverDate>
<prism:coverDisplayDate>14 August 2013</prism:coverDisplayDate>
<prism:doi>10.1371/journal.pone.0069841</prism:doi>
<citedby-count>45</citedby-count>
<affiliation>
<affilname>University Michigan Ann Arbor</affilname>
<affiliation-city>Ann Arbor</affiliation-city>
<affiliation-country>United States</affiliation-country>
</affiliation>
<pubmed-id>23967061</pubmed-id>
<prism:aggregationType>Journal</prism:aggregationType>
<subtype>ar</subtype>
<subtypeDescription>Article</subtypeDescription>
<article-number>e69841</article-number>
<source-id>10600153309</source-id>
</entry>
</search-results>
```

Figura 65: Ejemplo XML de respuesta de SCOPUS a través del navegador Firefox.

Fuente: Autor.

Para el consumo de la API de ELSEVIER en el presente proyecto específicamente para el consumo del recurso de búsquedas de SCOPUS a nivel de programación, basta con la concatenación de caracteres siguiendo la sintaxis para conformar la URL de búsqueda utilizando el código de la Figura 17:

```
function _getCitas($doi)
{
    $citas=array();
    $scopus_url = ELSEVIER_URL.$this->_scopusKey."&query=DOI(\"\".$doi.\"");
    $scopus_response = file_get_contents($scopus_url);
    $scopus_json = json_decode($scopus_response);
    if ($scopus_json->{'search-results'}->{'opensearch:totalResults'}==0){
        $citas+=array("scopus"=>0);
    } else{
        $citas+=array("scopus"=>intval($scopus_json->{'search-results'}->{'entry'}[0]->{'citedby-count'}));
    }
    .
    .
    .
    return array($citas,$altmetricMoreInfo);
}
```

Figura 66: Código para obtener información de ELSEVIER.

Fuente: Autor.

d. Consumo de Altmetric

El servicio de Altmetric se lo utiliza para obtener información de varios sitios web, en su mayoría redes sociales, accedendo a una sola URL. Altmetric recopila la información de varias redes sociales tanto de ocio como las consideradas para fines investigativos como por ejemplo Mendeley o CiteULike.

Similar a ELSEVIER, Altmetric hace uso de una API Key para obtener resultados; esta llave no es obligatoria al momento de realizar la petición de información, pero solo se podrá realizar una cantidad limitada de consultas (720 peticiones de información cada hora), es decir que Altmetric posee una forma menos restrictiva de brindar la información al usuario. Si se desea aumentar la cantidad de peticiones por hora, se puede comprar una o también existe la posibilidad de adquirir una API Key escribiendo un mail al grupo de Altmetric y solicitar una llave de autenticación para fines investigativos. En la Figura 67, se puede observar el correo que se envió y su respuesta de cómo se la ha conseguido la API Key para el presente proyecto.

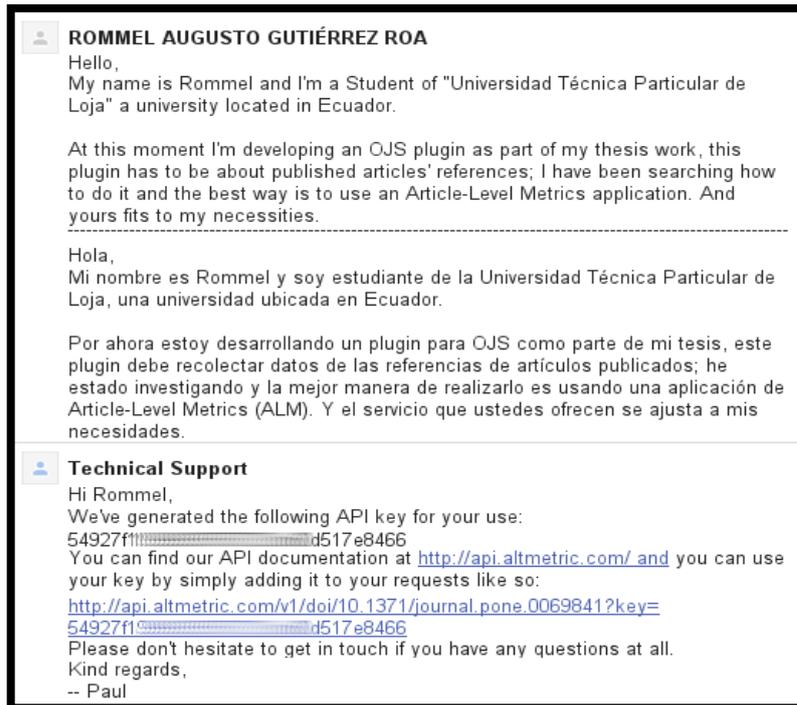


Figura 67: Obtención de la API Key de Altmetric.
Fuente: Autor.

Una vez obtenida la API Key, probamos en un navegador cómo funciona el API de Altmetric; para ello debemos formar la URL que se parece en algo a la de SCOPUS, pero con algunas variantes de sintaxis: API URL + DOI + API Key.

<http://api.altmetric.com/v1/doi/10.1371/journal.pone.0069841?key=APIKEYOBTENIDO>

Al navegar hacia la URL formada, obtendremos un documento en formato JSON el cual nos dará toda la información que existe para dicho DOI.

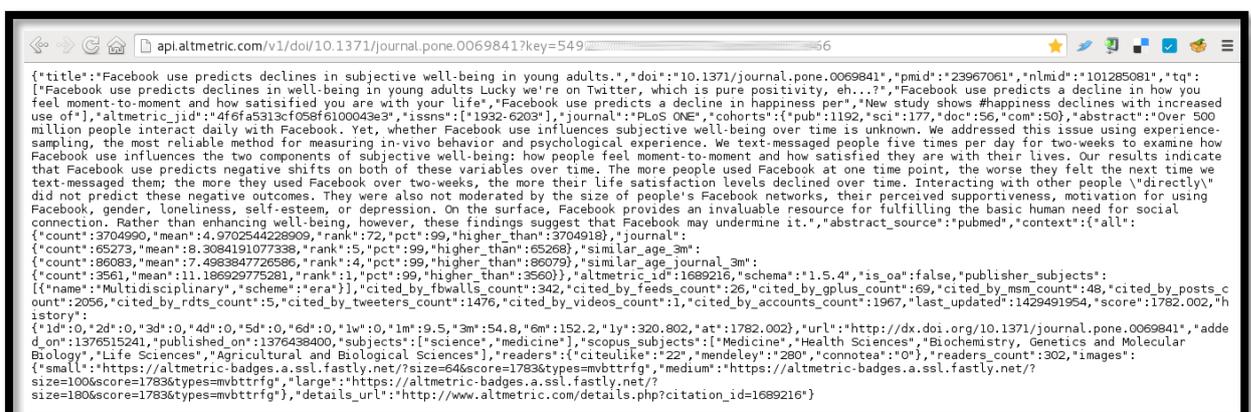
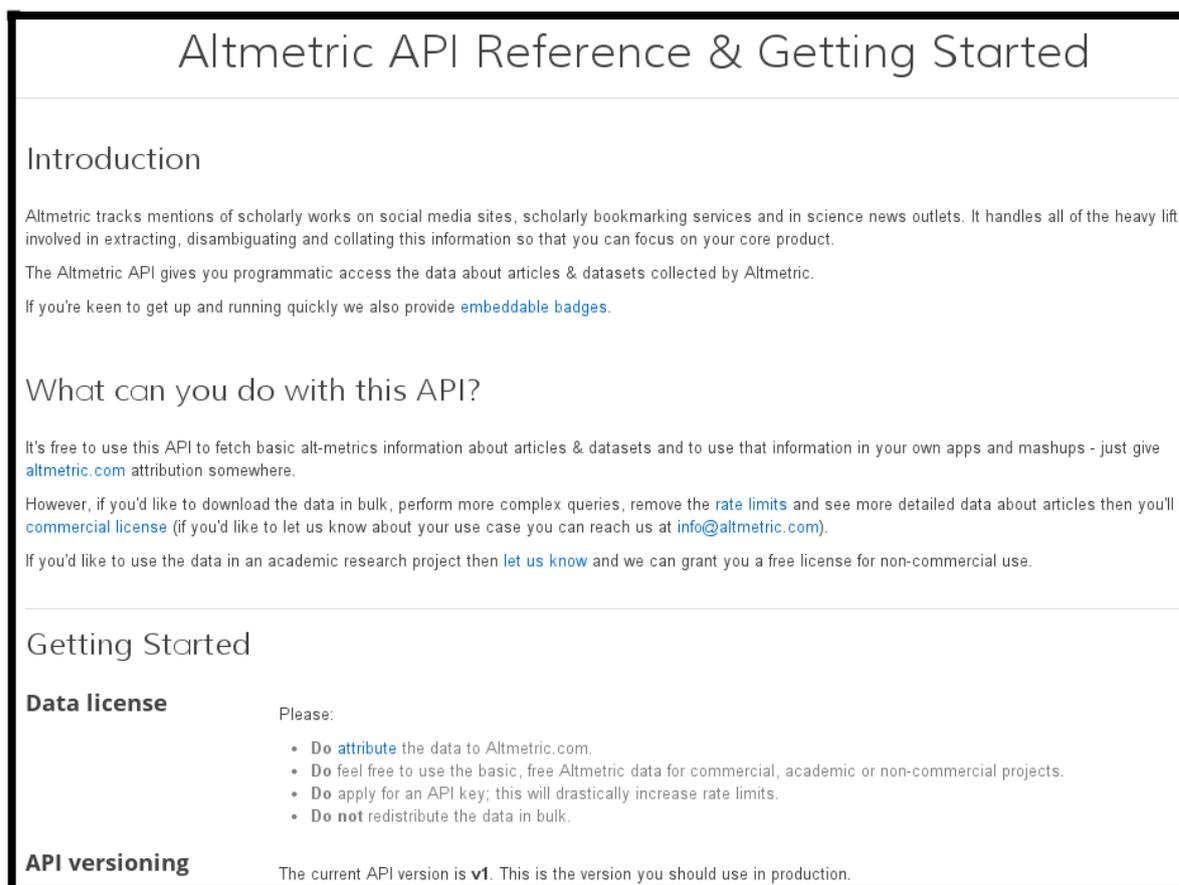


Figura 68: Información del DOI mediante el API Altmetric.
Fuente: <http://api.altmetric.com/>

Altmetric también posee una documentación detallada para poder hacer uso de su API, dicha documentación es mucho más corta que la de ELSEVIER y mucho más fácil de entender. Para poder acceder a este documento en línea, debemos acceder a la siguiente URL:

<http://api.altmetric.com/index.html>



The screenshot shows the 'Altmetric API Reference & Getting Started' page. It includes sections for 'Introduction', 'What can you do with this API?', and 'Getting Started'. The 'Getting Started' section contains sub-sections for 'Data license' and 'API versioning'. The 'Data license' section lists four bullet points: 'Do attribute the data to Altmetric.com', 'Do feel free to use the basic, free Altmetric data for commercial, academic or non-commercial projects', 'Do apply for an API key; this will drastically increase rate limits', and 'Do not redistribute the data in bulk'. The 'API versioning' section states 'The current API version is v1. This is the version you should use in production.'

Figura 69: Documentación del API de Altmetric.
Fuente: Autor.

En este sitio podemos observar que las búsquedas se pueden realizar únicamente a través del DOI y que el resultado que el API nos retorna está en formato JSON. Además indica que la sintaxis para el correcto uso del API es mediante como está a continuación:



The screenshot shows a browser address bar with the URL: `api.altmetric.com/v1/doi/10.1371/journal.pone.0069841?key=549`. The number 56 is visible at the end of the address bar.

Figura 70: URL formada para la consulta del DOI en el API de Altmetric.
Fuente: Autor.

Una vez que la petición se realiza correctamente, Altmetric retorna la información del DOI buscado en formato JSON el cual nos ofrece seis etiquetas con información relevante para enriquecer la información del artículo cuantitativamente.

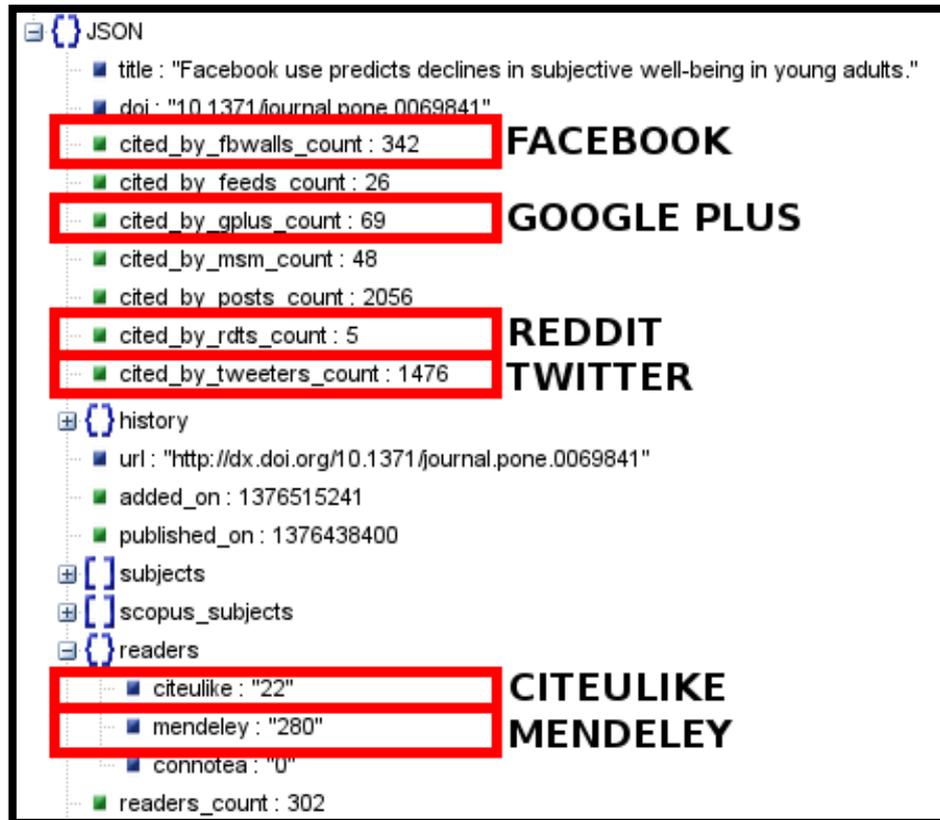


Figura 71: JSON de Altmetric

Fuente: Autor a través de <http://jsonviewer.stack.hu/>

E. Plugin. Manual de Usuario.

a. Instalación del plugin.

Para poder instalar el plugin, se debe acceder a la revista con un usuario que posea derechos de administrador. Una vez dentro, podremos observar las opciones que se detallan en la Figura 72.

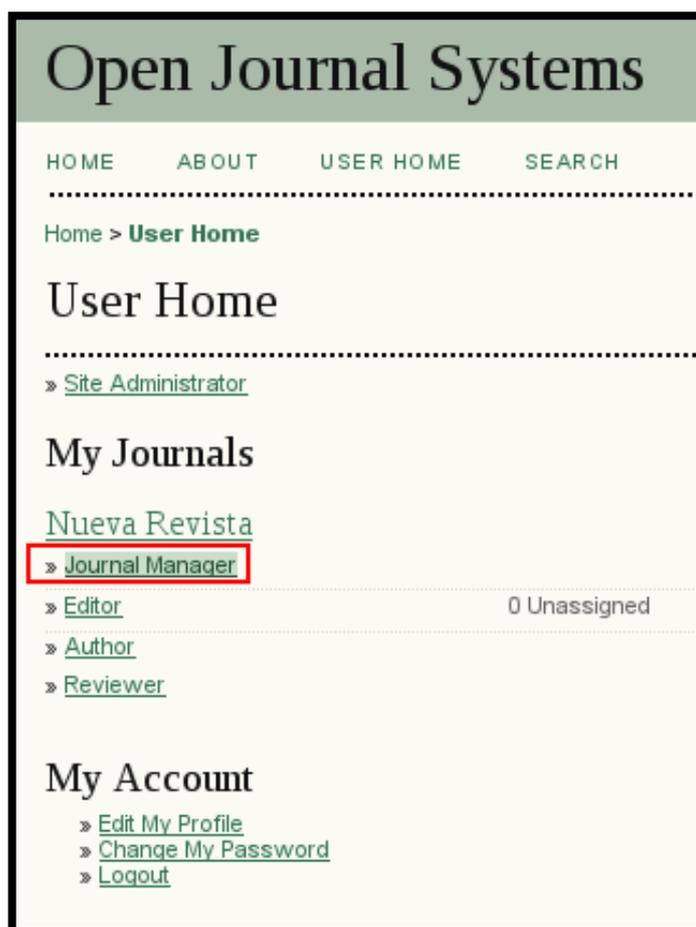


Figura 72: Ingreso al Administrador de la revista
Fuente: Autor.

De estas opciones, debemos elegir "Journal Manager". La cual nos llevará a un listado de páginas para la administración de nuestra revista. De este listado debemos seleccionar la opción "System Plugins" como se señala en la Figura 73.



Figura 73: Páginas de administración.
Fuente: Autor.

OJS tiene un sinnúmero de plugins, estos están organizados en grupos de acuerdo a la función que realizan. Para instalar nuestro plugin, seleccionamos la opción “Install A New Plugin” como se muestra en la Figura 74.

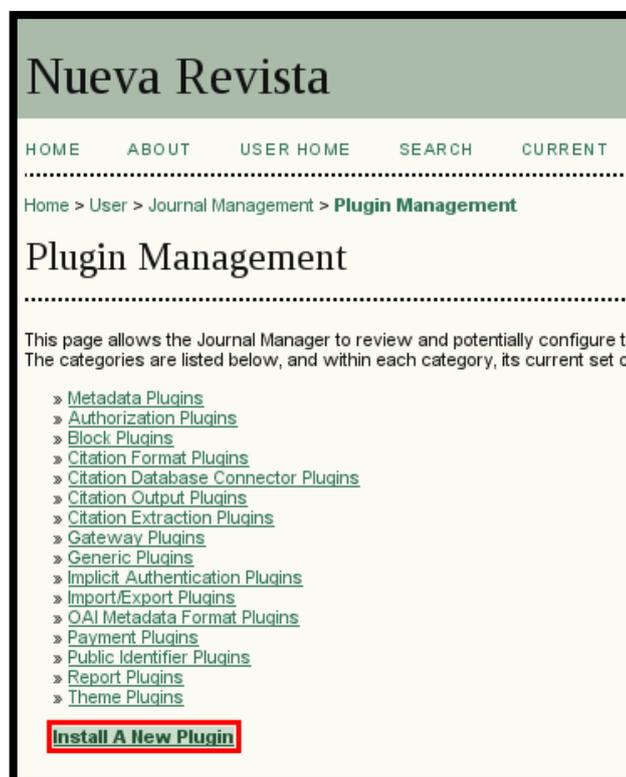


Figura 74: Administración de plugins.
Fuente: Autor.

La selección anterior, nos llevará a un formulario en el cual podremos elegir un archivo de nuestro computador. Este archivo debe ser un comprimido que tenga el plugin desarrollado. Dicho archivo comprimido debería estar con una extensión “tar.gz”.



The screenshot shows a web interface for 'Nueva Revista'. At the top, there is a navigation menu with links: HOME, ABOUT, USER HOME, SEARCH, CURRENT, and ARCHIVES. Below the menu is a breadcrumb trail: Home > User > Journal Management > Plugin Management > Install A New Plugin. The main heading is 'Install A New Plugin'. Below the heading, there is a message: 'This form allows you to upload and install a new plugin. Please ensure the plugin is compressed as a .tar.gz file.' At the bottom, there is a 'Select plugin file' label, a file input field containing 'utplMetrics.tar.gz', and a 'Continue' button.

Figura 75: Formulario de instalación de plugin.
Fuente: Autor.

Una vez seleccionado el archivo, hacemos click en el botón “continue” como observamos en la Figura 75, con el cual se empezará el proceso de instalación del plugin. Una vez culminado dicho proceso, se mostrará un mensaje el cual nos dará a conocer la versión del plugin instalado. En este caso la versión “1.0.0.0” como se puede ver en la Figura 76.



The screenshot shows the same web interface as Figure 75, but with a success message. The breadcrumb trail is: Home > User > Journal Management > Plugin Management > Install A New Plugin. The main heading is 'Install A New Plugin'. Below the heading, there is a message: '» Successfully installed version 1.0.0.0'. At the bottom, there is a 'Select plugin file' label, a file input field containing 'No se ha seleccionado ningún archivo.', and a 'Continue' button.

Figura 76: Mensaje de éxito al instalar plugin.
Fuente: Autor.



Figura 77: Administración de plugins, Plugins genéricos
Fuente: Autor.

Para configurar el nuevo plugin instalado, debemos regresar al menú de administración de plugins, y seleccionar la categoría “Generic Plugins” como se muestra en la Figura 77; esta selección nos llevará a un listado extenso de todos los plugins genéricos que nuestra revista posee, debemos ubicar el que lleva por nombre “utplMetrics Plugin” como se muestra en la Figura 78. Una vez ubicado el plugin, debemos seleccionar “Enable” para habilitar las funcionalidades de nuestro plugin.



Figura 78: Habilitación de plugin utplMetrics.
Fuente: Autor.



Figura 79: Confirmación de habilitado.
Fuente: Autor.

La habilitación del plugin, se verá al instante ya que este plugin fue diseñado para cambiar también la apariencia de la revista. A demás una notificación nos informará de la activación realizada, tal como se muestra en la Figura 79.



Figura 80: Configuración de utplMetrics.
Fuente: Autor.

Finalmente, deberemos ingresar la información de las llaves obtenida en el nexo "D", apartado "c", tal como se muestra en la Figura 81. Al guardar la información, otra notificación se mostrará indicando que se ha cambiado los datos con éxito.

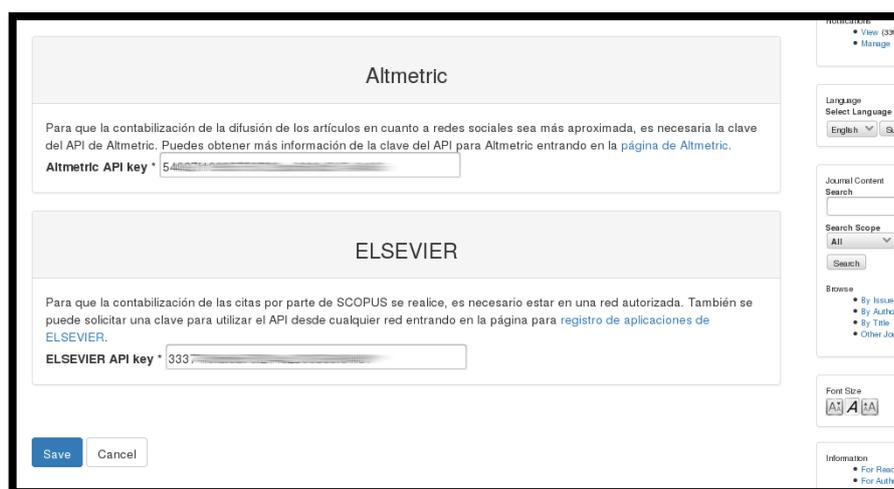


Figura 81: Notificación de configuración cambiada.
Fuente: Autor.

F. Marketing. Manual de Programador.

Para el desarrollo de esta herramienta aprovecharemos el API de ELSEVIER para poder acceder a los datos de los artículos que este posee. Pero esta ocasión utilizaremos un recurso diferente del API ya que esta vez no queremos obtener el número de veces citado, deseamos ver el resumen que se ha guardado para poder extraer el correo en caso que este exista. La documentación de las maneras de acceder al resumen del artículo las podemos encontrar en la dirección web: <http://api.elsevier.com/content/abstract> (Figura 82)

Abstract Retrieval API

This represents the interface to retrieve a Scopus abstract document. The full abstract has rich metadata representation, which includes links to [author](#) and [affiliation](#) profiles. The text of these abstracts are available for searching through use of the [Scopus Search API](#).

The native format of the abstract text is XML and while portions of it can be translated to JSON, the complexity of the mark-up prevents the full abstract from being accurately represented in JSON format.

The primary schema for the 'item' element of the full abstract can be found at <http://schema.elsevier.com/dtds/document/abstracts/xocs-ani512.xsd>. Other supporting schemas can be found at <http://schema.elsevier.com/dtds/document/abstracts/>.

The following chart shows the [Abstract Retrieval Views](#).

Summary

Resource	Method	Description
http://api.elsevier.com/content/abstract/scopus_id/{scopus_id}	GET	Abstract Retrieval API: This represents retrieval of a SCOPUS abstract by Scopus ID. Each abstract is assigned a unique Scopus identifier (which is also used to create the electronic identifier).
http://api.elsevier.com/content/abstract/eid/{eid}	GET	Abstract Retrieval API: This represents retrieval of a SCOPUS abstract by EID (Electronic Identifier). Each abstract is assigned a unique Scopus identifier which is also used to create the electronic identifier.
http://api.elsevier.com/content/abstract/doi/{doi}	GET	Abstract Retrieval API: This represents retrieval of a SCOPUS abstract by DOI (Document Object Identifier). Note that not every abstract contains a DOI. If only a DOI is available and it cannot be found using this resource, it may still be found by running a Boolean search against the SCIDIR Search API index and using the abstract link (if returned) in those search results.
http://api.elsevier.com/content/abstract/pii/{pii}	GET	Abstract Retrieval API: This represents retrieval of a SCOPUS abstract by PII (Publication Item Identifier). Note that not every abstract contains a PII. If only a PII is available and it cannot be found using this resource, it may still be found by running a Boolean search against the SCIDIR Search API index and using the abstract link (if returned) in those search results.
http://api.elsevier.com/content/abstract/pubmed_id/{pubmed_id}	GET	Abstract Retrieval API: This represents retrieval of a SCOPUS abstract by MEDLINE ID.
http://api.elsevier.com/content/abstract/pui/{pui}	GET	Abstract Retrieval API: This represents retrieval of a SCOPUS abstract by PUI.

Figura 82: Documentación, obtención del resumen del artículo mediante el API de ELSEVIER.

Fuente: documentación para desarrolladores de ELSEVIER.

Mediante el recurso de ELSEVIER mencionado, obtendremos una página HTML la cual deberemos analizar para determinar, mediante la técnica WebScraping, el patrón de etiquetas HTML que contendrán la información objetivo que en este caso son:

- Título del Artículo.
- Nombre del Autor.
- Correo del Autor.

b. Análisis de la estructura HTML del resumen del artículo.

El objetivo del análisis manual de la estructura HTML en WebScraping es encontrar particularidades para programar un patrón único que no se pueda repetir dentro del mismo código HTML y así obtener los datos dentro de una página web.

En una página web se pueden utilizar muchas etiquetas como “a”, “div”, “p”, entre otras más de una sola vez y aunque el estándar prohíbe repetir las etiquetas de títulos como “h1”, no limita a que se pueda repetir esta etiqueta. Por ello para identificar con una manera más precisa las etiquetas nos debemos guiar también por los atributos de las etiquetas como “class” que definen el estilo visual y la etiqueta “id” que es un identificador único.

En la Figura 29 se puede observar como el resumen de un artículo es presentado mediante el navegador, mientras que en la Figura 30 se muestra el código HTML interpretado por el navegador. Para el Web Scraping es importante identificar las etiquetas html que contienen la información a extraer, en este caso el título, autor y correo se encuentran al principio del resumen, y esto facilita la búsqueda manual de un patrón para posteriormente programar la extracción automática.

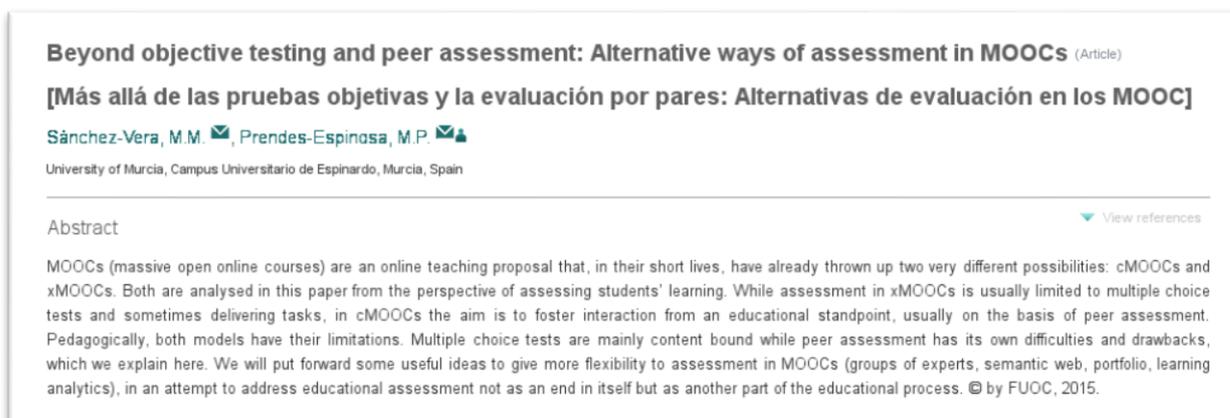


Figura 83: Ejemplo del resumen de un artículo

Fuente: Web de SCOPUS.



Figura 84: Ejemplo HTML del resumen de un artículo

Fuente: Web de SCOPUS

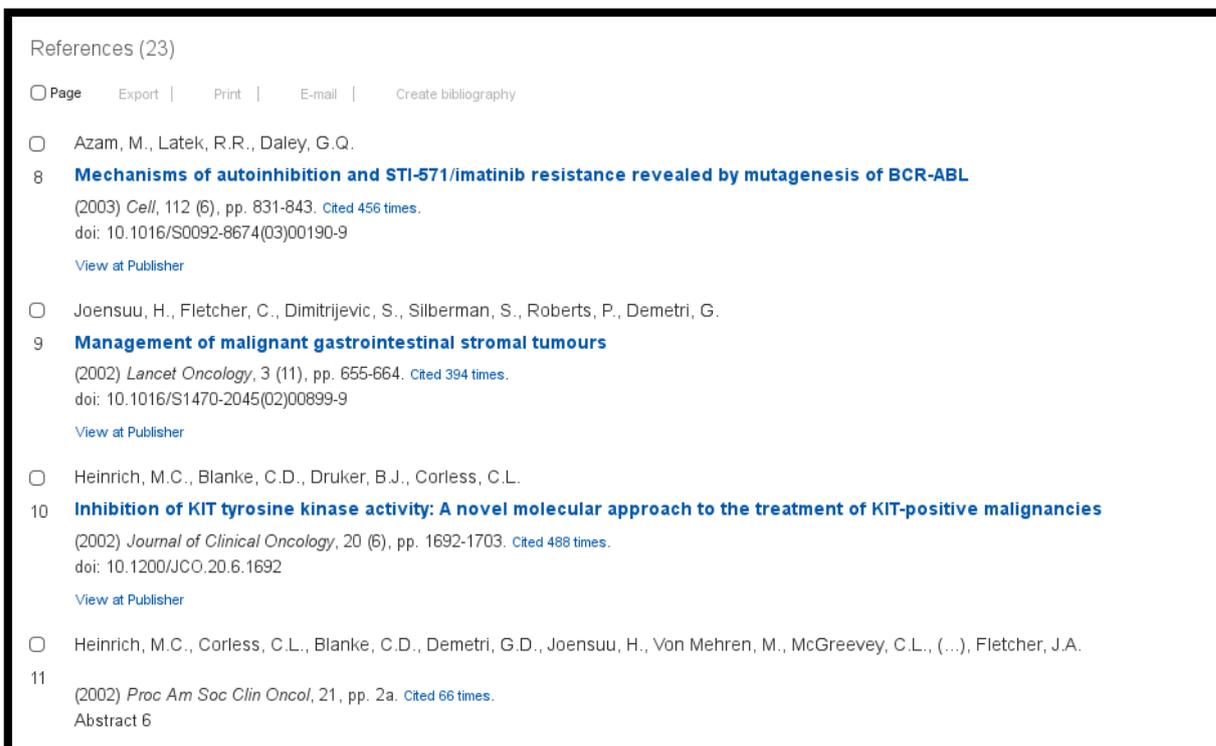
En el código HTML (Figura 30) se puede observar que el título del artículo se encuentra dentro de una etiqueta “h1” que es la etiqueta estandarizada para colocar el título principal de una página web, así mismo nos podemos dar cuenta que los nombres de los autores del

artículo se encuentran cada uno encerrado por las etiquetas “div” y “a” que son etiquetas comúnmente utilizadas para agrupar y crear enlaces a otras páginas respectivamente.

Una vez encontradas las etiquetas generales por decirlo así, debemos hacer que el patrón sea más específico utilizando los atributos “class” de las etiquetas. Mediante esta etiqueta se puede identificar que el correo electrónico de un autor tiene el valor “correspondenceEmail” para dicha etiqueta que en la Figura 29 está representado por la imagen de un sobre y que está agrupado dentro de una etiqueta “div” con el nombre del autor al que el correo pertenece; también se ha establecido que el título del artículo posee el valor “txtTitle” para el atributo class de la etiqueta “h1”.

De la misma manera debemos analizar el código HTML de las referencias del artículo (Figura 31) para extraer una lista de los artículos referidos que poseen DOI para obtener de estos la base de correos objetivo.

La identificación del patrón es un poco más fácil ya que se va a extraer una lista que en la página de SCOPUS está agrupada dentro de una etiqueta HTML “div” y cada ítem de la lista de referencias dentro de otra etiqueta “div” con el valor “referencesBlk” para el atributo “class”.



References (23)

Page Export Print E-mail Create bibliography

Azam, M., Latek, R.R., Daley, G.Q.
8 **Mechanisms of autoinhibition and STI-571/imatinib resistance revealed by mutagenesis of BCR-ABL**
(2003) *Cell*, 112 (6), pp. 831-843. Cited 456 times.
doi: 10.1016/S0092-8674(03)00190-9
[View at Publisher](#)

Joensuu, H., Fletcher, C., Dimitrijevic, S., Silberman, S., Roberts, P., Demetri, G.
9 **Management of malignant gastrointestinal stromal tumours**
(2002) *Lancet Oncology*, 3 (11), pp. 655-664. Cited 394 times.
doi: 10.1016/S1470-2045(02)00899-9
[View at Publisher](#)

Heinrich, M.C., Blanke, C.D., Druker, B.J., Corless, C.L.
10 **Inhibition of KIT tyrosine kinase activity: A novel molecular approach to the treatment of KIT-positive malignancies**
(2002) *Journal of Clinical Oncology*, 20 (6), pp. 1692-1703. Cited 488 times.
doi: 10.1200/JCO.20.6.1692
[View at Publisher](#)

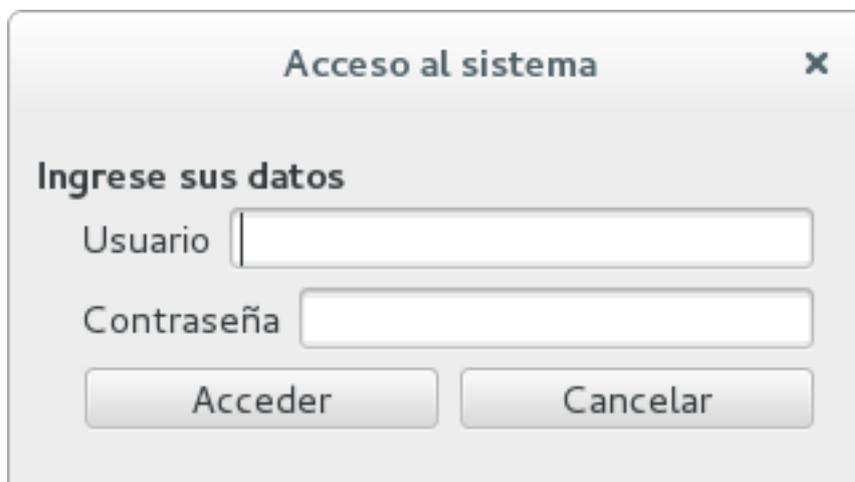
Heinrich, M.C., Corless, C.L., Blanke, C.D., Demetri, G.D., Joensuu, H., Von Mehren, M., McGreevey, C.L., (...), Fletcher, J.A.
11
(2002) *Proc Am Soc Clin Oncol*, 21, pp. 2a. Cited 66 times.
Abstract 6

Figura 85: Referencias de un artículo en SCOPUS

Fuente: Web de SCOPUS

G. Marketing. Manual de Usuario.

La primera pantalla (Figura 86) permitirá ingresar al sistema para extraer los correos. Con la clave y usuario por defecto que son: “admin”, “admin” llenar el formulario y hacer click en el botón aceptar.

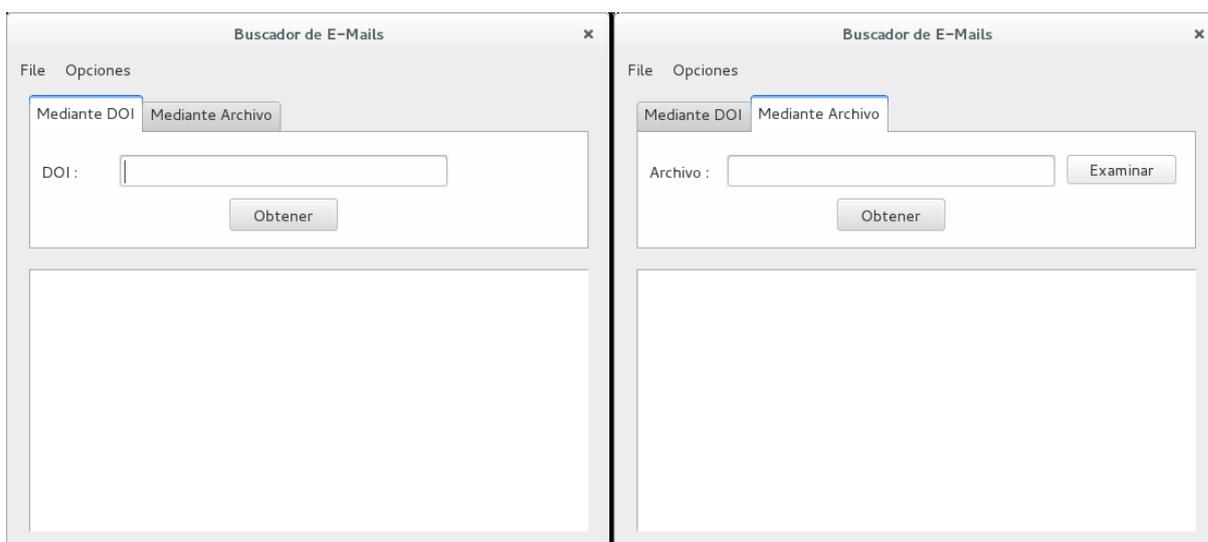


La imagen muestra una ventana de diálogo titulada "Acceso al sistema" con un botón de cerrar (X) en la esquina superior derecha. El título de la ventana es "Acceso al sistema". Debajo del título, el texto "Ingrese sus datos" indica que se trata de un formulario de inicio de sesión. Hay dos campos de entrada de texto: "Usuario" y "Contraseña". Debajo de los campos, hay dos botones: "Acceder" y "Cancelar".

Figura 86: ventana de ingreso a la herramienta de marketing.

Fuente: Autor.

Si los datos ingresados son correctos la aplicación mostrará la siguiente ventana en la cual el usuario podrá enviar a recuperar los correos de un artículo o de una lista de artículos mediante su DOI. En la Figura 87 podemos observar en la parte izquierda la interfaz para ingresar un solo DOI a procesar; y en la parte derecha la interfaz para añadir un archivo de texto con la lista de DOI de la cual deseamos obtener los correos.



La imagen muestra dos pantallas de la interfaz de usuario, ambas tituladas "Buscador de E-Mails". La pantalla de la izquierda tiene una pestaña "Mediante DOI" seleccionada y un campo de entrada etiquetado "DOI:" con un botón "Obtener" debajo. La pantalla de la derecha tiene una pestaña "Mediante Archivo" seleccionada y un campo de entrada etiquetado "Archivo:" con un botón "Examinar" a la derecha y un botón "Obtener" debajo. Ambas pantallas tienen una barra de menú con "File" y "Opciones" y un área de lista vacía en la parte inferior.

Figura 87: formas de ingresar DOI para procesar.

Fuente: Autor.

La Figura 88 muestra un ejemplo de un archivo de texto que contiene una lista de 13 DOI con la cual el programa puede trabajar para recuperar los correos de las referencias por cada artículo. La lista de DOI únicamente deberá contener el prefijo y el sufijo, sin anteponer el servicio de resolución de DOI de lo contrario la aplicación tomaría el DOI como inexistente.

```
10.7238/rusc.v12i1.2031
10.7238/rusc.v12i1.1947
10.7238/rusc.v12i1.2078
10.7238/rusc.v12i1.1944
10.7238/rusc.v12i1.1887
10.7238/rusc.v12i1.2475
10.7238/rusc.v12i1.2253
10.7238/rusc.v12i1.2185
10.7238/rusc.v12i1.2260
10.7238/rusc.v12i1.2262
10.7238/rusc.v12i1.2258
10.7238/rusc.v12i1.2296
10.7238/rusc.v12i1.2446
```

Figura 88: Ejemplo del archivo con lista de DOI.

Fuente: Autor.

Luego de introducir ya sea el DOI o el Archivo con los DOI, debemos hacer click en el botón “Obtener” el cual comenzará el proceso de extracción de datos, los botones y cuadros de texto se deshabilitarán mientras se realiza la obtención de datos. Al finalizar el proceso, aparecerá un botón al final de la ventana con el cual podremos guardar la información extraída en un archivo CSV tal como podemos apreciar en la Figura 89.



Figura 89: Fin de la ejecución de la herramienta de marketing.

Fuente: Autor.