

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-27

**DESARROLLO DE UNA PLATAFORMA DE MIGRACIÓN DE ARTÍCULOS
PUBLICADOS EN GRUPLAC A SCIENTIFIC JOURNAL RANKINGS (SJR)**

Alexander Sepúlveda Arenas

John Fredy Urrego Urrego

Robinson Alonso Román Hurtado

Ingeniería de Sistemas

Juan Sebastián Botero

INSTITUTO TECNOLÓGICO METROPOLITANO

16 Enero 2017

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

RESUMEN

Actualmente para los investigadores científicos existe la necesidad de tener una plataforma que permita realizar una búsqueda de información duplicada entre las diferentes publicaciones científicas aprobadas por pares académicos, hoy esta práctica debe realizarse manualmente para poder comparar individualmente los documentos en la información contenida, esta revisión demanda una gran cantidad de tiempo y desgaste por parte de los investigadores, además también existe la posibilidad de cometer errores y no poder identificar publicaciones que contengan campos con la información repetida al momento de ejecutar una comparación entre ellas.

De la misma manera, en Colombia no se cuenta con una herramienta que cumpla con los requerimientos necesarios y que pueda identificar duplicidad en los campos de las publicaciones científicas que se requieren validar como por ejemplo: ISSN (International Standard Serial Number), línea de investigación, grupo de investigación, nombre del proyecto, universidad, publicación y fecha de publicación.

Este proyecto se centró en la obtención rápida y efectiva de información mediante un cargue masivo de archivos referentes a artículos y revistas científicas desde la plataforma o base de datos de Colciencias. El proyecto permitió visualizar la información que estaba albergada en archivos html (Hyper Text Markup Language), extraer de forma rápida los campos requeridos de esta y facilitar al consultante obtener una respuesta veraz y factible de la búsqueda que se estaba realizando.

Los archivos se encuentran aprobados por los pares académicos y almacenados en la plataforma GrupLAC (Grupo Latinoamerica y del Caribe) y son migrados hacia una base de datos donde se indexan en el motor de almacenamiento mysql, se valida la información necesaria y posteriormente se puede comparar con el ranking de publicaciones que se descargan en un archivo de Excel desde el SJR (Scientific Journal Rankings).

Palabras clave: migración, SJRSCImago, cienciaometría, Scopus, GrupLAC, bibliometría, CvLAC.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

RECONOCIMIENTOS

Este trabajo de tesis ha sido una gran experiencia para evaluar todos los conocimientos adquiridos en una gran institución como lo es ITM, la cual cuenta con un gran plantel de docentes. Quiero agradecer especialmente a Juan Sebastián Botero, quién fue un gran apoyo en la culminación de este proceso, donde cada momento vivido fue simplemente único y cada oportunidad de corregir un error sin importar la cantidad de ellos durante el día a día, fue para lograr un aprendizaje futuro.

Agradezco a Dios y en especial a mi familia que fue el pilar más importante para no rendirme en los momentos difíciles, gracias a ellos una de mis metas está cumplida.

Alexander Sepúlveda Arenas

Agradezco a todos los que hicieron parte de esta trascendente etapa de mi vida, fue un ciclo de aprendizaje y experiencias al lado de profesores y compañeros de clases; gente valiosa de la que me llevo buenos recuerdos y mucha gratitud. Agradezco a mi familia que nunca dejó de creer en mí.

Agradezco a mi esposa por su paciencia y apoyo incondicional en esta etapa de mi vida.

John Fredy Urrego Urrego

Quiero primero agradecer a Dios por haberme permitido realizar sin interrupción esta hermosa carrera de la cual quedan muchos aprendizajes.

Agradezco a mi esposa e hijo quienes han permitido que pueda dedicar gran parte de mi tiempo para el desarrollo de este trabajo.

Agradezco a los profesores que con tanto esfuerzo y dedicación compartieron con nosotros todo su conocimiento para que lo apliquemos en nuestras vidas laborales y/o personales.

Agradezco a Juan Sebastián Botero, quien nos presentó la oportunidad de realizar este trabajo de grado y de quien pudimos aprender mucho durante nuestro proceso de formación.

Por último quiero agradecer a mis jefes y compañeros de trabajo en especial a Stephens Quiroz quien nos compartió su conocimiento y tiempo

Robinson Alonso Román.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

ACRÓNIMOS

CvLAC Curriculum vitae de Latinoamérica y del Caribe

SJR Scimago Journal & Country Rank

CTI Ciencia, Tecnología e Innovación

GrupLac Hoja de vida de grupos de investigación colombianos del sistema nacional en ciencia

IEEE Asociación mundial de ingenieros dedicada a la estandarización y el desarrollo en áreas técnicas.

SCIENCE Instituto para la información científica.

MEDLINE Base de datos de bibliografía médica más amplia que existe en el mundo.

SCOPUS Es la mayor base de resúmenes y referencias bibliográficas de literatura científica revisada por pares

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

TABLA DE CONTENIDO

Contenido

1. INTRODUCCIÓN	8
1.1. DESCRIPCIÓN DEL PROYECTO.....	8
1.2. JUSTIFICACIÓN.....	9
1.3. PLANTEAMIENTO DEL PROBLEMA.....	10
1.4. OBJETIVOS	10
1.4.1. General.....	10
1.4.2. Específicos.....	11
2. MARCO TEÓRICO	12
2.1. ESTADO DEL ARTE.....	12
3. METODOLOGÍA	20
3.1. METODOLOGÍA PROPUESTA.....	20
3.1.1. Pre-análisis	21
3.1.2. Análisis.....	22
3.1.3. Realización de diagramas de secuencia.....	23
3.1.4. Diseño:	23
3.1.5. Codificación:	24
3.1.6. Pruebas:	25
3.1.7. Resultados:	25
4 RESULTADOS Y DISCUSIÓN	27
5 CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO	32
REFERENCIAS	33
Referencias	33
APÉNDICE	35
ANEXOS	54

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

LISTA DE FIGURAS.

<i>Figura 1. Diagrama de casos de uso del sistema.....</i>	22
<i>Figura 2. Diseño de reportes</i>	24
<i>Figura 3. Grupos de investigación ITM.....</i>	27
<i>Figura 4. Publicaciones de grupos de investigación</i>	28
<i>Figura 5. Publicaciones duplicadas de grupos de investigación.....</i>	28
<i>Figura 6. Publicaciones grupo de investigación en historia, trabajo, sociedad y cultura</i>	29
<i>Figura 7. Menú de la plataforma SIA</i>	54
<i>Figura 8. Menú de carga de archivos html.....</i>	54
<i>Figura 9. Listado de universidades.....</i>	55
<i>Figura 10. Listado de grupos de investigación.....</i>	55
<i>Figura 11. Listado de publicaciones.....</i>	56
<i>Figura 12. Listado de publicaciones duplicadas</i>	56
<i>Figura 13. Botón de reinicio de base de datos.....</i>	56

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

LISTA DE ANEXOS

Anexo 1. Manual SIA Html.....	54
-------------------------------	----

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

1. INTRODUCCIÓN

1.1. DESCRIPCIÓN DEL PROYECTO

Actualmente, el nivel de desarrollo de los países está determinado por su capacidad de asimilar, producir y utilizar conocimiento científico en beneficio de la comunidad. Temas relacionados directamente con la ciencia, la tecnología, la innovación han favorecido notablemente la solución de los problemas que aquejan a la sociedad mediante la aplicación efectiva del conocimiento.

Para lograr la aplicación del conocimiento científico, este debe ser primero difundido, esta actividad fundamental de difusión del conocimiento a la sociedad en general, se realiza mediante la publicación de los resultados de las investigaciones científicas. Sin embargo, no siempre es fácil almacenar, clasificar y relacionar toda la información científica global garantizando tener los últimos avances para un tema en específico

Como herramienta para la difusión del conocimiento hoy se cuenta con la plataforma SJR (SCImago Journal Ranking), esta es un portal que incluye las revistas y los indicadores científicos de países desarrollados a partir de la información contenida en el Scopus base de datos (Elsevier), esta plataforma SJR permite realizar mediciones, según el número de publicaciones de revistas que sean aprobadas, y con ellas genera unos indicadores por universidades y países; dichos indicadores se pueden utilizar para evaluar y analizar el ranking científico en el que se encuentra un país [1].

A nivel local se tiene la base de datos de Colciencias donde se almacena todo el conocimiento científico generado en el país. Sin embargo, a diferencia de otras plataformas, esta no cuenta con la posibilidad de realizar análisis y relacionar temas científicos además de la revisión y verificación de duplicidad de trabajos de investigación

Actualmente la práctica de relación y revisión de información es realizada de forma manual y con una alta demanda en tiempo por parte de los investigadores. Además que su característica manual hace que se acarren errores, por ejemplo por revisión de información repetida

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Se pretenden entonces crear una nueva plataforma para optimizar el tiempo dedicado a la revisión de artículos científicos de acuerdo a campos requeridos como el ISSN, nombre de la universidad, nombre del proyecto y su fecha de publicación, además de minimizar el margen de error con la información ya existente. La plataforma permitirá realizar una comparación masiva de artículos y al finalizar su proceso de ejecución podrá mostrar además si hay publicaciones repetidas para una investigación ya aprobada.

En este proyecto se realiza una migración de datos de publicaciones científicas aprobadas por los pares académicos hacia una base de datos en mysql para poder ser consultada por investigadores, universidades y otras personas que requieran esta información para futuras investigaciones. Esta migración se realiza desde unos archivos html utilizando herramientas de uso libre como PHP (Hypertext Preprocessor), un lenguaje interpretado de bajo consumo de hardware y de fácil acceso desde cualquier equipo de cómputo por medio de una conexión a internet o por medio de un servidor local haciendo uso de un navegador web como: Chrome, Mozilla, Zafari; etc. También utiliza el motor de base de datos mysql que es una herramienta de uso libre, que ofrece la posibilidad de almacenar, editar, borrar y consultar grandes cantidades de información por medio de consultas

1.2. JUSTIFICACIÓN

Este proyecto se desarrolló debido a la necesidad que tienen los investigadores, universidades y estudiantes de realizar consultas de contenido científico de forma ágil y consistente, optimizando tiempos de ejecución en su revisión, teniendo en cuenta campos de validación como el ISSN, que es asignado a una publicación de manera única y no debe estar duplicado en la base de datos de Colciencias; además de otros campos como: línea de investigación, grupo de investigación, nombre del proyecto, universidad, publicación y fecha de publicación.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Es de vital importancia garantizar que todo el conocimiento científico que se desarrolla día a día y que se plasma en las publicaciones científicas sea conocido por estudiantes, investigadores u otras personas que requieren información actualizada y confiable para el desarrollo posterior de algún proyecto y es aún más importante que entidades como Colciencias tengan la información de revistas, artículos o documentos científicos ya aprobados por los pares académicos con un filtro de validación en el que se pueda encontrar si existe información duplicada.

1.3. PLANTEAMIENTO DEL PROBLEMA

Actualmente Colciencias cuenta con información muy importante dentro de sus bases de datos, dicha información por estar contenida en archivos html no puede ser visualizada fácilmente por personas que buscan información específica o referencias como base para iniciar una investigación o adelantar proyectos de grado. Por lo anterior sería de gran ayuda para Colciencias contar con una plataforma que le permita migrar toda su información a una base de datos y que desde allí pueda ser utilizada fácilmente para entre otros, verificar duplicidad en documentos y revisar el ranking de las universidades Colombianas que desarrollan artículos de investigación o científicos.

1.4. OBJETIVOS

1.4.1. General.

Diseñar e implementar una plataforma de software que permita validar la categoría de artículos de investigación usando la información reportada en GrupLAC con el ranking de SCImago Journal Report.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

1.4.2. Específicos.

- Desarrollar una plataforma que permita extraer información de artículos científicos de archivos html GrupLAC.
- Implementar una interfaz que permita conectar la información de GrupLAC con el SJR y visualizar la información de ranking de los artículos científicos.
- Generar informes de ranking para todos los artículos incluidos en la base de datos y poder ser comparados con otras plataformas.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

2. MARCO TEÓRICO

2.1. ESTADO DEL ARTE.

En la historia se encuentra que Eugene Garfield, en 1951 inicia a trabajar en la Universidad Johns Hopkins sobre un nuevo proyecto de indización automática en la biblioteca de la facultad de medicina; dicho proyecto llamado "Proyecto Welch", fue financiado directamente por la biblioteca nacional de medicina y su finalidad era tener allí reunida toda la información posible sobre medicina. Garfield finalizó este proyecto en 1953 y pudo obtener la información de cómo era manejada la estructura lingüística de las revisiones y como se manejaba el método tradicional de indización.

Mientras Garfield realizaba su trabajo encontró algo muy curioso y fue la relación de cómo eran referenciadas las ideas en la construcción de un nuevo artículo científico. Al concluir la idea de lo anterior, pensó que si se seguía la vida de un artículo, se podría encontrar también el desarrollo de su idea principal, cuál era su uso y cuáles serían sus posibles implementaciones. Entonces, Garfield resumió que la mejor forma de conocer toda la vida de un artículo sería realizando una indización por citas.

Es común recurrir entonces a disciplinas métricas como la bibliometría y la cienciometría, que sirven de gran ayuda no sólo para describir y analizar las propiedades y características del proceso de producción científica, tecnológica y del que hacer investigativo en general, sino como insumo para optimizar la toma de decisiones, y dirigir las al bienestar de la sociedad.

El método de la bibliometría se ha aplicado históricamente para buscar la relación existente entre las investigaciones, para encontrar el número de citas que se han proporcionado y para ello se aplican unas leyes que permiten determinar cuál es el impacto en dichas citas por medio de gráficos estadísticos.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

La Bibliometría sirve para poder identificar cual ha sido el crecimiento en el campo de investigación científica y cuáles son los países o regiones que puntúan en referencias con la utilización de sus investigaciones.

La cienciometría se encarga del estudio de la ciencia como problema multidimensional donde puede reunir investigadores, libros y conocimiento; este método puede medir el nivel de desarrollo y aporte a la ciencia en su parte cuantitativa en cuanto a la productividad y creatividad de los investigadores. Tiene sus orígenes y auge en 1977 con la creación de Scientometrics 15, que es originaria de Hungría; luego aparece la cienciometría moderna en los trabajos de Derek J. de Solla Price y Eugene Garfield.

Se debe observar cuál es la funcionalidad que cumple la cienciometría para poder entender lo que se puede llevar a cabo con ella y con qué fines se utiliza en el ámbito de la producción científica; para ello se muestra cómo se aplica hoy en día la cienciometría, como por ejemplo el SJR-SCImago que está dedicada al análisis de la información, la representación y la recuperación por medio de técnicas de visualización y además realiza un análisis con base en la información obtenida para relacionarla

En conclusión, la bibliometría y la cienciometría son dos caminos estrechamente relacionados con respecto a las mediciones que se realizan a las publicaciones científicas y permiten obtener unos indicadores de números de citas hechas por autores en documentos de investigación y revistas científicas.

Las plataformas del CvLAC (Curriculum Vitae de Latinoamérica y el Caribe) y el SJR (SCImago Journal Rankings), realizan mediciones de prestigio para las publicaciones de artículos y revistas de investigación en el ámbito científico; estas mediciones se basan en la idea de que “no todas las citas fueron creadas de igual manera”.

El papel que juega cada una de estas plataformas (CvLAC, SJR, Scopus, GrupLAC) en la publicación y modificación de la información es usado para el análisis científico de artículos y revistas para tener como “factor de impacto” criterios de la evaluación, la calidad de revistas y artículos científicos.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

El término "factor de impacto" fue concebido por Eugene Garfield en 1955 para ayudar a los bibliotecarios a identificar las revistas más prestigiosas en función del número de citas. El primer ranking de revistas por "factor de impacto" fue publicado en 1972 [2]. Estas revistas de factor de impacto y algunas otras mediciones bibliométricas son muy utilizadas en varios países para conocer cuál es el prestigio que tienen sus investigaciones.

En conjunto el CvLAC y el GrupLAC, son las herramientas que constantemente se están actualizando en la gran biblioteca de datos científicos que tiene Colciencias; en estas dos plataformas los investigadores, innovadores, estudiantes y científicos que deseen, pueden ingresar para inscribir su hoja de vida y proyectos de investigación.

Existen además en el mundo, entidades que se encargan de realizar publicaciones de las diferentes áreas del conocimiento, entre ellas se encuentran la IEEE (Institute of Electrical and Electronics Engineers), SCIENCE, MEDLINE, etc. Existe también Elsevier, la cual creó un sistema de bases de datos llamada Scopus que permite consultar información como: perfiles de autor, universidades o realizar un rastreo de citas de revistas, documentos o artículos científicos que ya han sido aprobados por los pares de cada una de las áreas de investigación, además de que se puede conocer el factor de impacto de las publicaciones.

El factor de impacto se mide también con la información que está almacenada en una base de datos en polibuscador, que se llama JCR (Journal Citation Reports).

Los requerimientos que se deben cumplir para estar incluido en un ranking de revista de factor de impacto son: Criterio de evaluación de la relación entre la calidad de las revistas y el factor de impacto, la evaluación de la calidad de las publicaciones, el período de evaluación para el cálculo del índice de impacto, el cálculo del índice de impacto, la comparación entre revistas de diferentes campos científicos, la generalización del factor de impacto de una revista a cada uno de los artículos publicados en ella, la relación entre el factor de impacto y la dificultad de publicar, el idioma en que publican las revistas y el cómo mejorar el factor de impacto de una revista [3].

Las nuevas tecnologías de la información han visto la necesidad de crear plataformas en las cuales se pueda encontrar o consultar información acerca de las investigaciones científicas

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

que ya han sido aprobadas por los pares académicos para su continuidad en la búsqueda de soluciones a problemas o necesidades que se encuentran en la sociedad.

Colciencias ente que es regulado directamente por la presidencia de la república y que en sus bases de datos contiene la biblioteca nacional en Colombia de artículos o documentos científicos ya aprobados requiere, que antes de publicarse cualquier investigación, esta sea verificada por los investigadores científicos y así evitar duplicidad en la información; los investigadores podrán contar con una aplicación en la cual tendrán la facilidad de comprobar si en las revistas o documentos científicos que ya han sido aprobados por los pares académicos a diferentes universidades existe duplicidad de información; esta información será llevada a una base de datos y podrá ser consultada en cualquier momento por instituciones académicas, estudiantes o cualquier persona que la requiera para su conocimiento.

El área encargada de revisar los comportamientos en la investigación y realizar análisis de todas las publicaciones científicas como libros, revistas o artículos es la bibliometría; está en una parte tiene relación con la cienciometría que aplica métodos matemáticos y estadísticos a toda la literatura de carácter científico y a los autores que la producen con el objetivo de estudiar y analizar la actividad científica, para ello se ayuda de leyes bibliométricas basadas en el comportamiento estadístico regular que a lo largo del tiempo han mostrado los diferentes elementos que forman parte de la ciencia. Los instrumentos utilizados para medir los aspectos de este fenómeno social son los indicadores bibliométricos, medidas que proporcionan información sobre los resultados de la actividad científica en cualquiera de sus manifestaciones [4].

Eugene Garfield, fundó el Instituto para la información científica, este es usado frecuentemente para el análisis bibliométrico. Garfield en sus investigaciones sobre la bibliometría propone siempre que en toda investigación que requiera realizar citas estas sean hechas correctamente y que por medio de las correctas citas se dé el reconocimiento a los autores iniciales que ya se desgastaron un buen tiempo para iniciar con la búsqueda de información en un tema específico. Los investigadores recurren a la

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

citación de seguimiento para encontrar los artículos más prestigiosos para un tema en particular y para ver con qué frecuencia se citan sus propios artículos publicados. Durante años los investigadores que buscaban este tipo de información sólo tenía un recurso para consultar: Web of Science de Thomson Scientific. En 2004 surgieron dos competidores Scopus de Elsevier y Google Scholar de Google.

La cienciaometría que se encarga del estudio de la ciencia como problema multidimensional donde puede reunir investigadores, libros y conocimiento; realiza consultas sobre la importancia que tienen para las investigaciones científicas el funcionamiento de las plataformas del CvLAC (Curriculum Vitae de Latinoamérica y el Caribe) que es la base de datos donde quedan registradas todas las hojas de vida que se inscriben en Colciencias [5]. Este instrumento CvLAC es una herramienta que fue creada por Colciencias para:

- *El registro permanente de hojas de vida de investigadores, innovadores, académicos y expertos del SNCyTI. El registro de hojas de vida de candidatos a ingresar al servicio de información de pares evaluadores reconocidos del SNCyTI. Las convocatorias para reconocimiento y medición de Grupos Colombianos de Investigación Científica o Tecnológica [5].*
- *Convocatorias de becas para formación de recurso humano, apoyo a programas doctorales a nivel nacional y jóvenes investigadores [5].*

La información consignada en CvLAC, es individual y debe ser diligenciada personalmente por cada investigador quien realiza la actualización a través de una clave única de acceso al sistema [5].

GrupLAC “Es un producto del CNPq (Consejo Nacional de Pesquisas del Brasil) adaptado especialmente para Colombia, es usado para el diligenciamiento y actualización de los datos de cada una de las hojas de vida inscritas en el CvLAC de Colciencias y así estas puedan ser posteriormente utilizadas por los investigadores, estudiantes, científicos o grupos de investigación para ser consultados en el SNCyT” [6].

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

El SJR (SCImago Journal & Country Rank) es usado para la publicación de artículos y revistas de investigación en el ámbito científico; es desarrollado por SCImago Lab y utiliza tecnología de Scopus. El SJR es una medición de prestigio basada en la idea de que no todas las citas fueron creadas de igual manera; donde las citas de un artículo se consideran una medida del reconocimiento científico de estudio que ha recibido, y por lo tanto un indicador de su valor y el impacto en el campo científico; el SJR también tiene relación con la plataforma de SCIMAGO INSTITUTIONS RANKINGS [7].

El SCImago Journal & Country Rank es un portal que incluye las revistas y los indicadores científicos de países desarrollados a partir de la información contenida en el Scopus; estos indicadores son utilizados para evaluar y analizar los dominios científicos. Esta plataforma toma su nombre del indicador SCImago Journal Ranking (SJR); que permite tener la visibilidad de las revistas incluidas en el Scopus base de datos a partir de 1996, el SJR calcula también su propio factor de impacto basándose en la información que está incluida en el Scopus, de la empresa Elsevier.

Scopus realiza una serie de comparaciones con todos los documentos que se encuentran inscritos en las bases de datos SJR y Elsevier, esta arroja un resultado con el número de citas que se han presentado con cada autor o artículo [8].

Dentro de todas las bases de datos que existen hasta el momento, Scopus es la base de datos bibliográfica más grande del mundo que contiene resúmenes y citas de artículos de revistas científicas. Cubre aproximadamente 18.000 títulos de más de 5.000 editores internacionales, incluyendo la cobertura de 16.500 revistas revisadas por pares de las áreas de ciencias, tecnología, medicina y ciencias sociales, incluyendo artes y humanidades; es la base de datos en la que todas las universidades se apoyan para la búsqueda de información requerida para el complemento de las investigaciones y es en la que se puede encontrar información con mayor eficacia y confiabilidad [9].

En varios casos se realizan comparaciones entre algunas bases de datos para encontrar la que más se pueda acomodar a la necesidad de consultas. Según Laura M. Faltes, en su artículo *“Google Scholar, Scirus, and the Scholarly Search Revolution”*, habla de una

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

comparación entre dos bases de datos, Google Scholar y la comparan con Elsevier's Scirus; para encontrar cuál de las dos puede ofrecer mejor información para los investigadores [10]. Esta comparación arroja finalmente un resultado en que se da a conocer cuál de estas dos bases de datos tiene mayor y mejor información para ser utilizada por los investigadores

Hoy en día, Google Scholar (<http://scholar.google.com>) ofrece un nuevo método para localizar los artículos potencialmente relevantes sobre un tema determinado mediante la identificación de los artículos posteriores que citan un artículo publicado anteriormente. Una característica importante de Google Scholar es que los investigadores pueden realizar rastreos de interconexiones entre los autores que citan artículos sobre el mismo tema y determinar la frecuencia con que otras personas citan un artículo en específico, ya que tiene un “citado por” característica. Este estudio también compara los recuentos de citas proporcionadas por Web of Science y Google académico para los artículos en el campo de la "Webometrics". Por último, concluye que Google académico ofrece una alternativa libre o complemento de otros índices de citas [11].

Dentro de sus actualizaciones y mejoras en la información almacenada en las bases de datos, Google académico pretende incluir "artículos revisados por pares, tesis, libros, resúmenes e informes técnicos de editoriales académicas, sociedades profesionales, depósitos de impresiones preliminares y universidades, así como artículos académicos disponibles a través de la Web. Algunas colecciones prominentes incluyen ACM, Annual Reviews, arXiv, Blackwell, IEEE, Ingenta, Instituto de Física, Sistema de datos de Astrofísica de la NASA, PubMed, Nature Publishing Group, RePEc (Cuadernos de Investigación en Economía), Springer y Wiley Interscience, aunque no todos en su totalidad. Muchos sitios web de las universidades y las organizaciones sin fines de lucro están incluidas, pero sólo los documentos que parecen como artículos de revistas académicas [12].

Se conoce también, el portal del SCImago Journal & country Rank, que también permite realizar consultas de información científica, Esta herramienta permite tener datos

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

estadísticos de los artículos más consultados; la información para esta es tomada desde la base de datos Scopus de Elsevier.

El SJR es una fuente con acceso abierto que ofrece valiosa información para la evaluación métrica de la ciencia; a esta plataforma se le da su nombre por el indicador SCImago Journal Rank (SJR); esta plataforma utiliza para su búsqueda y ordenamiento el algoritmo de Google más conocido como Google Pagerank; este permite visualizar la información contenida en la base de datos Scopus la cual es alimentada desde 1996.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3. METODOLOGÍA

3.1. METODOLOGÍA PROPUESTA.

Se inició con la etapa de toma de requerimientos para poder identificar cuál era el problema que se venía presentando, encontrando que el problema más complejo y que debía tenerse como prioridad era la duplicidad que presentaba la información, adicionalmente se debía buscar una forma ágil y practica para poder realizar estas búsquedas.

Entonces se desarrolló una plataforma de migración de artículos publicados en GrupLAC para ser comparados con Scientific Journal Ranking, implementada utilizando una base de datos mysql donde son indexados archivos que Colciencias tienen en formato html.

Con el desarrollo de esta plataforma fue incluida una interfaz que permite realizar el cargue masivo de la información que se va a procesar, esta información debe ser cargada desde una ubicación especifica en la que se tengan almacenados los archivos. Una vez cargada la información, la plataforma inicia su procesamiento para la obtención de los resultados.

Desde los archivos publicados en GrupLAC de Colciencias se extrajeron de la base de datos campos puntuales como: ISSN, línea de investigación, grupo de investigación, nombre del proyecto, universidad, publicación y fecha de publicación. Con la extracción de estos campos la plataforma genera unos informes en pantalla donde se puede visualizar toda la información que fue procesada y se verifica cuál de estas se repite.

En los informes generados se puede identificar cuáles son los grupos de investigación, líneas de investigación y nombre de universidades que han realizado publicaciones y puede ser comparada con el ranking que genera el SJR.

La interfaz del proyecto se desarrolló usando el framework bootstrap; este es de código abierto para el diseño de sitios y aplicaciones web y se ajusta a la necesidad presentada. Este framework está compuesto por los lenguajes html y css, además de extensiones de Javascript. La interfaz además contiene plantillas de diseño, formas, botones, navegación y otros componentes gráficos

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Bootstrap es compatible con las últimas versiones de navegadores web como chrome, Mozilla Firefox, Opera, Internet Explorer y Safari. En el desarrollo de esta plataforma se utilizó la versión 3.3.7; esta versión es compatible con el diseño web adaptable, esto significa que el diseño de las páginas se ajusta dinámicamente al dispositivo que se utiliza ya sean computadores de escritorio, tabletas o teléfonos móviles.

Todo lo anterior se llevó a cabo mediante la ejecución de las fases de investigación sobre la Cienciometría, investigación sobre SJR-SCLmago, desarrollo del motor de búsqueda y modelamiento de la base de datos, teniendo en cuenta los siguientes pasos:

- Pre análisis
- Análisis
- Diseño
- Codificación
- Pruebas
- Resultados

3.1.1. Pre-análisis

- ***Recolección de requerimientos:*** En esta etapa se definió de manera detallada las necesidades que se pretenden cubrir con el desarrollo del proyecto.
- ***Realización de la narrativa:*** En esta etapa se creó un texto que contiene la descripción detallada del método de búsqueda y el cual llevará las tareas del sistema, el procesamiento de los datos, las decisiones del sistema y la entrada/salida de información.
- ***Realización de clases:*** En esta etapa se identificó, a partir de la narrativa, las clases potenciales para la ejecución del proyecto con sus respectivos atributos y operaciones que permitieran continuar con su desarrollo.
- ***Realización de Casos de uso:*** En este paso se describió la secuencia de eventos de los actores que intervienen en el método de búsqueda, teniendo en cuenta la

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

3.1.3. Realización de diagramas de secuencia.

En esta etapa se representa gráficamente la secuencia lógica de los casos de uso, los eventos generados por los actores externos, su orden y los eventos internos del sistema.

El sistema de migración de datos presenta un esquema general que se basa en los siguientes requerimientos funcionales:

- Hacer compatibles las plataformas de origen y destino de datos.
- Establecer la conexión con la fuente de datos
- Capturar la información del origen de datos.
- Establecer la conexión con el destino de los datos.
- Migrar la información. Validar la información migrada
- Generar el informe de consistencia de la migración

3.1.4. Diseño:

En esta parte se buscó describir y construir una estructura lógica que diera solución al sistema planteado en el análisis. Para ello se siguieron los siguientes pasos:

- **Revisión del análisis:** Se hizo un estudio al análisis para identificar posibles errores y realizar sus respectivas correcciones.
- **Realización del diseño de la base de datos:** En esta parte se estableció en detalle cómo será la estructura física, tablas, atributos, relaciones y formas de acceso de la información indexada.
- **Realización de diseño de reportes:** En esta fase se diseñó una pantalla que permite consultar y visualizar los resultados obtenidos cuando se realiza la migración de la información a la base de datos.

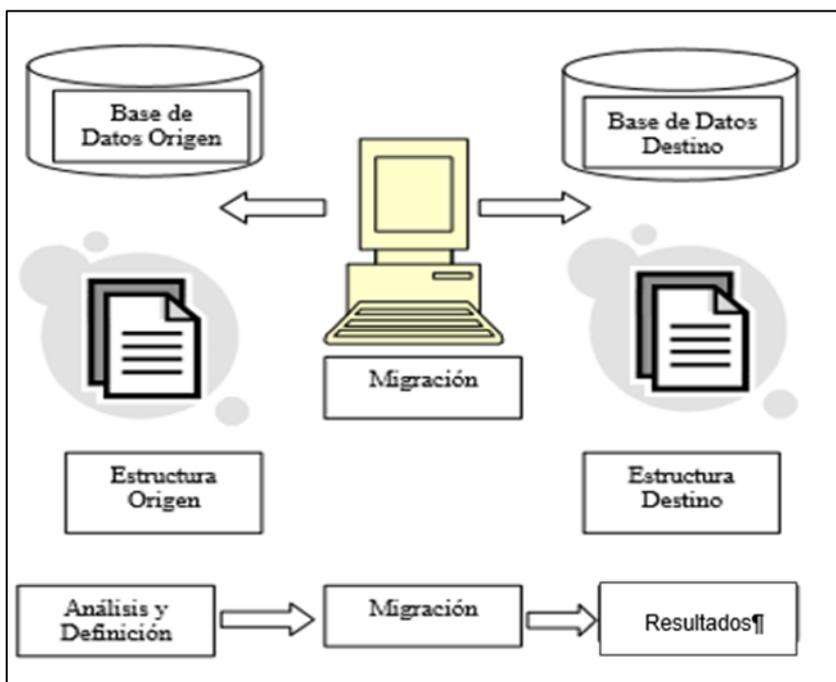


Figura 2. Diseño de reportes

3.1.5. Codificación:

Es la programación del aplicativo en el lenguaje más adecuado para el desarrollo del algoritmo. Implementando un método de búsqueda se realiza el proceso de migración de los datos sobre el “puente” construido. Una vez se ha establecido la conexión con el origen y el destino de los datos se procede a realizar la migración de los datos con las estructuras conectadas; el aplicativo migrador hace un reconocimiento de ambas estructuras y procede a generar un archivo interno con los datos origen convertidos a la estructura destino, tras lo cual se descargarán en la plataforma destino de la migración los datos migrados y validados estructuralmente.

Una vez superada la fase de análisis y diseño del sistema propuesto, se procede con el proceso de implementación o codificación y haciendo uso del método de búsqueda se construye una interfaz o módulo del migrador de bases de datos.

- **Selección del tipo de destino:** En esta sección se seleccionó el motor de base de datos al cual se va a conectar a la plataforma. Sobre dicho motor se va a ejecutar el proceso

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

de migración ubicando la ruta donde se encuentra la base de datos a migrar, para esto se escribe manualmente la ruta o se presiona el botón examinar el cual traerá una ventana de selección.

Para ejecutar la plataforma se selecciona el servidor y puerto de conexión. Se debe digitar la dirección y/o nombre del servidor donde se encuentra el motor a migrar, de igual manera se debe seleccionar el puerto que se está utilizando para la conexión al motor de la base de datos origen.

Luego de establecer la conexión se selecciona la base de datos que se va a migrar, para esto se utiliza el menú desplegable, el cual va a mostrar las bases de datos disponibles. Luego se presiona el botón conectar base de datos para establecer la conexión de esta base de datos y selección de la(s) tabla(s) a migrar

Para la ejecución se seleccionan las tablas que se van a migrar, así como la base de datos destino donde quedarán almacenados los datos ya migrados.

3.1.6. Pruebas:

En esta etapa se realizó la validación para saber si la información ingresa o no a la base de datos y se hicieron las respectivas correcciones mediante un plan de pruebas validando cada una de las opciones del aplicativo.

También durante el proceso de desarrollo del proyecto se fueron evaluando permanentemente los elementos que entraban a conformar el sistema propuesto.

3.1.7. Resultados:

Entre los eventos importantes que se presentaron durante el proceso de desarrollo del sistema, se tienen entonces:

La realización de un análisis para determinar las necesidades que deberá cubrir el proceso de migración de datos para generar un esquema metodológico unificado como el propuesto.

La construcción de los diagramas más representativos de búsqueda para modelar el sistema propuesto, como con: el diagrama de casos de uso y el diagrama de clases de los requerimientos funcionales más representativos del sistema.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

La construcción de un prototipo con el método de búsqueda del sistema a migrar de base de datos.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

4 RESULTADOS Y DISCUSIÓN

El proyecto titulado desarrollo de una plataforma de migración de artículos publicados en grupLAC a scientific journal rankings (SJR), muestra una propuesta de mejoramiento, en la búsqueda de información y revisión de publicaciones y revistas científicas aprobadas, para los grupos de investigación

Este proyecto permite de forma rápida y sencilla hacer la búsqueda de información de publicaciones, esta búsqueda se puede realizar cargando uno a uno los archivos en la plataforma o realizando un cargue masivo. De estos archivos se extrae la información contenida sobre universidades, grupos de investigación y sus publicaciones pudiendo verificar si estas últimas se encuentran duplicadas

En la siguiente figura podemos observar cómo se almacena la información extraída de los archivos html a la base de datos de mysql.

← T →		lin_intId	uni_intId	lin_strNombre
<input checked="" type="checkbox"/>		95	44	Grupo de Investigación en Integración...
<input checked="" type="checkbox"/>		96	44	Derecho y Territorio
<input type="checkbox"/>		97	44	Evolución del Relieve y los Suelos
<input type="checkbox"/>		98	44	Centro de estudios en Derecho Público "...
<input type="checkbox"/>		99	44	Estabilidad en Clases No Elementales
<input type="checkbox"/>		100	44	conflicto en el aula
<input type="checkbox"/>		101	44	HIDROFOBICIDAD DE SUELOS

Figura 3. Grupos de investigación ITM

Una vez cargados los archivos en el SIA (sistema de indexación de archivos) se aplica el método de búsqueda utilizado para la extracción de los datos. A continuación se muestran imágenes de como el SIA presenta la información extraída de los archivos html.

Inicio Cargar Archivo Universidades Grupo de Investigación Publicaciones Publicaciones Duplicadas Reiniciar DB

Listado de Publicaciones

Mostrar 10 registros Buscar: Buscar Registro

GRUPO INVESTIGACION	ID PUBLICACION	PUBLICACION	ISSN	Año
Epidemiología	3513	Conocimientos y prácticas frente a la exposición solar y tanzaje de cáncer de piel, en usuarios de una institución de salud prestadora de servicios de salud de Antioquia, 2000.	1657-0448	2001
Epidemiología	3514	Familias y escuelas de niños (as) con comportamientos agresivos y prosociales: Clave para su intervención.	1657-4400	2002
Epidemiología	3515	Clima escolar y comportamientos psicosociales en niños	0120-385X	2005
Epidemiología	3516	Caracterización del acceso a los servicios de salud en población infantil desplazada y receptora en asentamientos marginales en seis ciudades de Colombia, 2002 - 2003	0120-385X	2006
Epidemiología	3517	Conocimientos y prácticas del personal de salud acerca del diagnóstico de tuberculosis respiratoria en adultos, ciudad de Medellín, 2007	1909-1265	2008
Epidemiología	3518	Papel del personal de salud en el diagnóstico tardío de la tuberculosis pulmonar en adultos de Medellín, Colombia.	1539-5545	2010
Epidemiología	3519	Nutritional supplementation and physical work capacity in adolescent and young adult Guatemalans.	0022-3166	1995
Epidemiología	3520	Promoción de la salud	0120-5307	1998
Epidemiología	3521	Programas de rehabilitación Cardíaca en Colombia	0120-5307	1999
Epidemiología	3522	La actividad física en el ámbito de la Salud Pública	0120-385X	1998

Mostrando registros del 1 al 10 de un total de 4,308 registros Anterior 1 2 3 4 5 ... 431 Siguiente

Todos los derechos reservados a S I A - HTML - 2016 | Desarrollado por John Freddy Urrego, Robinson roman, Alexander Sepulveda. Versión 2016.05.01

Figura 4. Publicaciones de grupos de investigación

Listado de Publicaciones

Mostrar 10 registros Buscar: Modelo de Simulación Diná

GRUPO INVESTIGACION	ID PUBLICACION	PUBLICACION	ISSN	Año
GIP	141	Modelo de Simulación Dinámica para el Diagnóstico de la PYME	0121-750X	2005
Grupo de Investigación en Gestión Industrial GEGI	121	Modelo de Simulación Dinámica para el Diagnóstico de la PYME	0121-750X	2005

Mostrando registros del 1 al 2 de un total de 9.646 registros (filtrado de un total de 9.646 registros) Anterior 1 Siguiente

Listado de Publicaciones

Mostrar 10 registros Buscar: Las Empresas de prestaci

GRUPO INVESTIGACION	ID PUBLICACION	PUBLICACION	ISSN	Año
GIP	131	Las Empresas de prestaciñ de servicios y la determinaciñ de su capacidad de operaciones	0123-921X	2004
Grupo de Investigación en Gestión Industrial GEGI	113	Las Empresas de prestaciñ de servicios y la determinaciñ de su capacidad de operaciones	0123-921X	2004

Mostrando registros del 1 al 2 de un total de 9.646 registros (filtrado de un total de 9.646 registros) Anterior 1 Siguiente

Listado de Publicaciones

Mostrar 10 registros Buscar: Análisis de cambios de rég

GRUPO INVESTIGACION	ID PUBLICACION	PUBLICACION	ISSN	Año
RISE	5357	Análisis de cambios de régimen en series de tiempo no lineales utilizando modelos TAR	0120-2596	2004
Simulación y modelación matemática	1438	Análisis de cambios de régimen en series de tiempo no lineales utilizando modelos TAR	0120-2596	2004

Mostrando registros del 1 al 2 de un total de 9.646 registros (filtrado de un total de 9.646 registros) Anterior 1 Siguiente

Listado de Publicaciones

Mostrar 10 registros Buscar: Traqueostomía percutánea

GRUPO INVESTIGACION	ID PUBLICACION	PUBLICACION	ISSN	Año
Grupo de Investigación en Anestesia U de A	6984	Traqueostomía percutánea por dilatación	0120-3401	2000
Grupo de Investigación en Anestesia U de A	6980	Traqueostomía percutánea por dilatación	0120-3347	2002

Mostrando registros del 1 al 2 de un total de 9.646 registros (filtrado de un total de 9.646 registros) Anterior 1 Siguiente

Listado de Publicaciones

Mostrar 10 registros Buscar: Intubación fibróptica en pe

GRUPO INVESTIGACION	ID PUBLICACION	PUBLICACION	ISSN	Año
Grupo de Investigación en Anestesia U de A	6979	Intubación fibróptica en pediatría	0120-3401	2000
Grupo de Investigación en Anestesia U de A	7011	Intubación fibróptica en pediatría	0120-3401	2000
Grupo de Investigación en Anestesia U de A	7013	Intubación fibróptica en pediatría	0120-3347	2005

Mostrando registros del 1 al 3 de un total de 9.646 registros (filtrado de un total de 9.646 registros) Anterior 1 Siguiente

Figura 5. Publicaciones duplicadas de grupos de investigación

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

Los anteriores son ejemplos de los resultados que se obtienen sobre todas las publicaciones que sean cargadas en la plataforma y que están duplicadas. Se observa que son diferentes grupos de investigación y tiene la misma publicación realizada. Esta es la información que será comparada con el informe de ranking que genera la plataforma SJR.

S I A - HTML Inicio Cargar Archivo Universidades Grupo de Investigación Publicaciones Publicaciones Duplicadas Reiniciar DB

Listado de Publicaciones

Mostrar 10 registros Buscar:

GRUPO INVESTIGACION	ID PUBLICACION	PUBLICACION	ISSN	Año
Historia, Trabajo, Sociedad y Cultura	251	La broca del café en el Líbano. Impacto social y cultural en los años 90	0123-885X	2009
Historia, Trabajo, Sociedad y Cultura	252	Tendencias de la historia regional en Colombia. Problemas y perspectivas recientes	2145-132X	2011
Historia, Trabajo, Sociedad y Cultura	253	Sociedad, trabajo y población en Fredonia (Antioquia), 1830-1852. Una perspectiva demográfica y socio-económica según los censos y padrones	0121-8417	2010
Historia, Trabajo, Sociedad y Cultura	254	Estudios e historiografía del café en Colombia, 1970-2008. Una revisión crítica	0122-1450	2010
Historia, Trabajo, Sociedad y Cultura	255	Fundación del Líbano, (1849-1886)	0120-2456	2003
Historia, Trabajo, Sociedad y Cultura	256	Modernización del régimen laboral hacendatario. Un estudio sobre el impacto de reformas jurídicas en empresas cafeteras rurales cafeteras, Líbano-Tolima, 1959-1965	0122-1450	2003
Historia, Trabajo, Sociedad y Cultura	257	Exilio y Género, El caso de la mujer colombiana en Suecia	0348-145X	1998
Historia, Trabajo, Sociedad y Cultura	258	Formación de una hacienda cafetera: mecanismos de organización empresarial y relaciones administrativo-laborales. El caso de La Aurora (Líbano-Colombia), 1882-1907	0122-1450	1999
Historia, Trabajo, Sociedad y Cultura	259	El vestido nuevo y mi otro yo Colombi	1657-4303	2004
Historia, Trabajo, Sociedad y Cultura	260	Reseña: Cultivating Coffe. The Farmers of Carazo, Nicaragua, 1880-1930 de Julie Charlip.	0188-2503	2004

Mostrando registros del 251 al 260 de un total de 2.383 registros Anterior 1 ... 25 26 27 ... 239 Siguiente

Todos los derechos reservados a S I A - HTML - 2016 | Desarrollado por John Freddy urego, Robinson roman, Alexander Sepulveda. Versión 2016.05.01

Figura 6. Publicaciones grupo de investigación en historia, trabajo, sociedad y cultura

Esta nueva plataforma tiene como ventaja que se pueden procesar grandes cantidades de información contenida en los archivos html que son descargados desde la base de datos de Colciencias. Comparativamente con el tiempo que le toma a un investigador realizar la validación manual e individual de los archivos esta plataforma tiene unos tiempos de respuesta reducidos a segundos y cerrará una brecha que se tiene en el campo de la investigación.

Aún se presentan algunas limitaciones que no permitirán tener el funcionamiento de la plataforma al máximo y es que para conectarse directamente a la base de datos de Colciencias se requieren permisos exclusivos para extraer la información directamente y trabajar con la información más actualizada que exista en la base de datos.

El conocimiento generado a partir de la investigación del desarrollo tecnológico nacional debe tener un mayor impacto sobre el sector productivo y contribuir a la solución de las

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

problemáticas de la sociedad colombiana, para alcanzar este objetivo se han implementado una serie de estrategias y acciones para que la aplicación y apropiación del conocimiento sean más cercanos a las inquietudes y necesidades de la empresa y la sociedad en general.

Unas de las principales decisiones a tener en cuenta para la implementación del modelo de medición del conocimiento generado en el país que permita tener un mayor impacto de la investigación científica a nivel productivo son:

- Validar la vinculación de los investigadores y producciones registradas.
- Recomendar a las instituciones que para avalar sus grupos de investigaciones, desarrollo tecnológico o de innovación dispongan de un reglamento interno para dicho fin.
- Colciencias podrá solicitar a las instituciones o a las personas los soportes de nuevo conocimiento, desarrollo tecnológico e innovación y los de formación de capital humano en sus instituciones en la dependencia que el representante legal defina para tal fin.
- Los grupos de investigaciones, desarrollo tecnológico e innovación que se presente a la convocatoria de reconocimiento de grupos deberán tener en su poder evidencia física o digital de toda la producción que hayan inscrito en la plataforma con el fin de responder ante una eventual solicitud de Colciencias para su verificación.
- Este proceso es exclusivamente destinado a investigadores vinculados contractualmente a instituciones colombianas y residentes en el país.
- Los productos de único autor solo podrán ser vinculados por una sola vez, aun solo grupo de investigación.
- Para vincular un producto a un grupo deberá contar con la autorización respectiva del usuario de cada CvLAC.

Es preciso fortalecer un sistema científico que investigue en los distintos campos de la frontera de las ciencias, que permita el avance del conocimiento fundamental y que al mismo tiempo este orientado a promover la transferencia de ese conocimiento como un elemento esencial para la innovación y en últimas para el bienestar y sostenibilidad para todos los miembros de la sociedad; así como fortalecer un sistema creativo que permita

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

generar conocimiento sobre nuevas interpretaciones de la subjetividad del ser humano y sus relaciones con los demás.

Se ha podido observar como a lo largo del tiempo los investigadores y grupos de investigación científica han venido creciendo tanto en actualizaciones de proyectos existentes, como en nuevas propuestas y temas de interés para todo el círculo de investigadores científicos, estudiantes y universidades. Esta buena práctica ha hecho que Colciencias revise su infraestructura y plataformas para poder implementar modelos de medición de los cuales se pueda tener información importante para la solución de problemas en la población.

Para Colciencias fue muy importante la creación de la política de “Apoyo al fortalecimiento y consolidación de los grupos y centros de investigación del país”. A raíz de esta política se propusieron las definiciones para los grupos de investigación, líneas de investigación, investigadores y centros de investigación, también cuales serían los indicadores de medición. Para este fin Colciencias inicio con unas convocatorias para grupos y centros de investigación; de estas fueron citadas unas cinco en la década de los noventa con el fin de poder empezar con la recopilación de información de cuáles eran los grupos de investigación que para la fecha existían y así poder otorgar ayudas para su fortalecimiento.

Después de varios años, en 1998 se pudo crear el primer modelo de medición, después de este primer modelo y con el crecimiento que se ha presentado y que el tipo de medición debe cambiar, se han venido realizando varias modificaciones hasta la actualidad.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

5 CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO

-
- Este desarrollo permitirá a los grupos de investigación optimizar tiempo, pasando de búsquedas manuales a búsquedas automáticas de información de los trabajos, artículos y revistas ya aprobados para algún tipo de investigación.
 - Este desarrollo permitirá poder realizar la extracción de información de archivos html a una base de datos donde pueda ser analizados de manera más sistemática.
 - Para claridad de la información se recomienda crear una herramienta en la cual se puedan generar gráficos de medición en cuanto al ranking de cada universidad, sus grupos de investigación y su número de investigaciones publicadas.
 - Esta plataforma permitirá a futuro realizar mediciones de la información científica creada en el país y tomar decisiones sobre ella.
 - A futuro se puede tener en cuenta poder leer los archivos html directamente desde una URL (Uniform Resource Locator)
 - Se deberán gestionar permisos para poder acceder directamente a la base de datos de Colciencias.
 - Sería importante construir un webservice que nos permita interactuar directamente con la información disponible en la base de datos de Colciencias.
 - A futuro se debería generar reportes que permitan la opción de exportar los resultados a archivos de formatos EXCEL o PDF.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

REFERENCIAS

Referencias

- Colciencias. (s.f.). *CvLAC*. (Colciencias) Recuperado el 04 de 11 de 2015, de <http://www.colciencias.gov.co/glosario>
- Colciencias. (s.f.). *Software GrupLAC*. (Colciencias) Recuperado el 04 de 11 de 2015, de <http://www.colciencias.gov.co/glosario>
- Elsevier. (2016). *sobre Scopus*. (Elsevier) Recuperado el 07 de 02 de 2016, de <https://www.elsevier.com/solutions/scopus>
- Falter, L. (2005). Google Scholar, Scirus, and the Scholarly Search Revolution. <http://connection.ebscohost.com/c/articles/16014815/google-scholar-scirus-scholarly-search-revolution>, 13(2), 43.
- Garfield, E. (2003). *The meaning of the Impact Factor*. Recuperado el 07 de Febrero de 2016, de <http://garfield.library.upenn.edu/papers/meaningofif2003.pdf>
- Garfield, E. (04 de Enero de 2006). *The History and Meaning of the Journal Impact Factor*. Recuperado el 07 de Febrero de 2016, de <http://jama.jamanetwork.com/article.aspx?articleid=202114>
- Hill, K. (2007). Using Scopus to Analyze Citations. *The Carleston Advisor*, 9(2), 59-61.
- Meyer, E. T. (s.f.). *What is Bibliometrics and Scientometrics?* (TIDSR: Toolkit for the Impact of Digitised Scholarly Resources) Recuperado el 04 de Noviembre de 2015, de <http://microsites.oii.ox.ac.uk/tidsr/kb/48/what-bibliometrics-and-scientometrics>
- Noruzi, A. (2005). Google Scholar : the new generation of citation indexes. *e-prints in library & information science*, 55(4), 170-180.

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

R. Notess, G. (Agosto de 2005). *Scholarly Web Searching: Google Scholar and Scirus*. (Information Today, Inc.) Recuperado el 04 de Noviembre de 2015, de <http://www.infoday.com/Online/jul05/OnTheNet.shtml>

SCImago Journal & Country Rank. (04 de 07 de 2012). *scimagojr*. (Scimago Lab) Recuperado el 02 de 11 de 2015, de <http://www.scimagojr.com/>

Scimago, L. (s.f.). *SCImago Journal & Country Rank*. Recuperado el 04 de Noviembre de 2015, de <http://www.scimagojr.com/aboutus.php>

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

APÉNDICE

Apéndice A: Código fuente SIA

```
//-----index.php-----//
<?php include_once 'theme/header.php'; ?>
<?php include_once 'theme/menu.php'; ?>
<div class="container-fluid">
<div class="row">
<div class="col-sm-1 col-sm-offset-1 col-md-12 col-md-offset-0 main">
<p class="text-center">Pagina en construcci&oacute;n</p>
</div>
</div>
</div>
<?php include_once 'theme/footer.php'; ?>
<!-- Bootstrap core JavaScript
===== -->
<!-- Placed at the end of the document so the pages load faster -->
<script
src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.3/jquery.min.js"></script>
<script>window.jQuery || document.write('<script
src="../../assets/js/vendor/jquery.min.js"></script>')</script>
<script src="bootstrap-3.3.6/dist/js/bootstrap.min.js"></script>
<!-- Just to make our placeholder images work. Don't actually copy the next line! -->
<script src="bootstrap-3.3.6/assets/js/vendor/holder.min.js"></script>
<!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
<script src="bootstrap-3.3.6/assets/js/ie10-viewport-bug-workaround.js"></script>
<script src="bootstrap-3.3.6/datatables/dataTables.bootstrap.min.js"></script>
<script src="bootstrap-3.3.6/datatables/jquery.dataTables.min.js"></script>
<script src="http://ajax.googleapis.com/ajax/libs/jquery/1.7/jquery.js"></script>
<script src="js/jquery.form.js"></script>
<script src="js/bar.js"></script>
```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

</body>

</html>

//-----cargar_archivo.php-----//

<?php include_once 'theme/header.php'; ?>

<?php include_once 'theme/menu.php'; ?>

<div class="wrapper">

<div class="content-wrapper" style="background:#FFFFFF;">

<div class="container-fluid">

<div class="row">

<div class="col-sm-1 col-sm-offset-1 col-md-12 col-md-offset-0 main">

<div class="panel panel-primary">

<div class="panel-heading text-center">Agregar Archivos

html</div>

<div class="panel-body">

<form class="form-horizontal" action="upload.php" method="post" enctype="">

<div class="form-group">

<div class="col-xs-5">

<input type="file" name="archivo[]" id="archivo[]" multiple="multiple" required>

</div>

<div class="col-xs-1">

<button type="submit" class="btn btn-primary"></

span> Cargar Archivo</button>

</div>

</div>

</form>

<hr>

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<div class="progress">
<div class="progress-bar progress-bar-striped active bar" role="progressbar">
<span class="percent">0%</span>
</div>
</div>
<div id="status" class="text-center"></div>
</div>
<?php include_once 'theme/footer.php'; ?>
<!-- Bootstrap core JavaScript
===== -->
<!-- Placed at the end of the document so the pages load faster -->
<script
src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.3/jquery.min.js"></script>
<script>window.jQuery || document.write('<script
src="../../assets/js/vendor/jquery.min.js"></script>')</script>
<script src="bootstrap-3.3.6/dist/js/bootstrap.min.js"></script>
<!-- Just to make our placeholder images work. Don't actually copy the next line! -->
<script src="bootstrap-3.3.6/assets/js/vendor/holder.min.js"></script>
<!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
<script src="bootstrap-3.3.6/assets/js/ie10-viewport-bug-workaround.js"></script>
<script src="bootstrap-3.3.6/datatables/dataTables.bootstrap.min.js"></script>
<script src="bootstrap-3.3.6/datatables/jquery.dataTables.min.js"></script>
<script src="http://ajax.googleapis.com/ajax/libs/jquery/1.7/jquery.js"></script>
<script src="js/jquery.form.js"></script>

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<script src="js/bar.js"></script>
</body>
</html>

//-----upload.php-----//

<?php
include_once('config/Conexion.php');
$content = mb_convert_encoding($content, 'HTML-ENTITIES', 'UTF-8');
mysql_set_charset('utf8');
//comprobamos que sea una petición ajax
if (!empty($_SERVER['HTTP_X_REQUESTED_WITH']) &&
strtolower($_SERVER['HTTP_X_REQUESTED_WITH']) ==
'xmlhttprequest') {
# definimos la carpeta destino
$carpetaDestino = "files/";
# si hay algun archivo que subir
if ($_FILES["archivo"]["name"][0]) {
# recorremos todos los archivos que se han subido
for ($i = 0; $i < count($_FILES["archivo"]["name"]); $i++) {
# si es un formato html
if ($_FILES["archivo"]["type"][$i] == "text/html") {
# si existe la carpeta o se ha creado
if (file_exists($carpetaDestino) || @mkdir($carpetaDestino)) {
$origen = $_FILES["archivo"]["tmp_name"][$i];
$destino = $carpetaDestino . $_FILES["archivo"]["name"][$i];
# movemos el archivo
if (@move_uploaded_file($origen, $destino)) {
$directorio = opendir("files/");
while ($documentosHTML = readdir($directorio)) { //obtenemos un archivo y luego otro
sucesivamente

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

if ($documentosHTML == "." || $documentosHTML == ".." || $documentosHTML ==
"Parent Directory") {
//do nothing
} else {
$doc = new DOMDocument();
libxml_use_internal_errors(true);
$doc->loadHTMLFile("files/" . $documentosHTML);
$xpath = new DOMXPath($doc);
$selements = $xpath->query("*/span[@class='titulo1']");
$grupoinvestigacion = $selements->item(0)->nodeValue;
$selements = $xpath->query("*/table/tr/td[@class='celdaEncabezado']");
$universidad = "";
$producciones = array();
if (!is_null($selements)) {
foreach ($selements as $element) {
if (strpos($element->nodeValue, 'Instituciones') !== FALSE) {
$parenttr = $element->parentNode;
$nexttr = $parenttr->nextSibling;
$universidad = trim($nexttr->nodeValue);
$universidad = substr($universidad, strpos($universidad, ".-") + 3, strpos($universidad, "
-") - 3);
}
if (strpos($element->nodeValue, 'Produc') !== FALSE) {
$parenttr = $element->parentNode;
$nexttr = $parenttr->nextSibling;
if ($nexttr) {
while ($nexttr) {
$text = trim($nexttr->nodeValue);
$inicio = strpos($text, ":") + 1;
$fin = strpos($text, "\n");

```

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

$produccion["nombre"] = substr($text, $inicio, $fin);
$inicio = strpos($text, "ISSN:") + 6;
$fin = strpos(substr($text, $inicio), ",");
$produccion["issn"] = substr($text, $inicio, $fin);
$inicio = strpos($text, "ISSN:") + 17;
$fin = strpos(substr($text, $inicio), "vol:");
$produccion["fecha"] = substr($text, $inicio, $fin);
if (strlen($produccion["nombre"]) > 0 && strlen($produccion["issn"]) > 0 &&
strlen($produccion["fecha"]) > 0) {
$producciones[] = $produccion;
}
$nexttr = $nexttr->nextSibling;
}
}
}
}
}
$conn = new Conexion();
$conn->conectar();
if ($universidad != "") {
$universidad = htmlentities($universidad, ENT_QUOTES, 'UTF-8'); //convertimos el
string a utf-8 para
evitar problemas con los asientos
$sqluniversidad = "INSERT INTO tbl_universidad (uni_intId,uni_strNombre) VALUES
('," . $universidad
. "')";
//echo "<BR>SQL: ".$sqluniversidad;
$conn->consultar($sqluniversidad);
}
if ($grupoinvestigacion != "") {

```

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

$sqlconsultauniversidad = "SELECT MAX(uni_intId) AS uni_intId FROM
tbl_universidad";
$respon = $conn->consultar($sqlconsultauniversidad);
$row = mysql_fetch_row($respon);
$id_universidad = trim($row[0]);
$grupoinvestigacion = htmlentities($grupoinvestigacion, ENT_QUOTES, 'UTF-8');
//convertimos el
string a utf-8 para evitar problemas con los asentos
$sqlgrupoinvestigacion = "INSERT INTO tbl_lineainvestigacion (lin_intId, uni_intId,
lin_strNombre)
VALUES ('', '' . $id_universidad . '', '' . trim($grupoinvestigacion) . '')";
$conn->consultar($sqlgrupoinvestigacion);
}
if ($producciones != "") {
$sqlconsultalinea = "SELECT MAX(lin_intId) AS lin_intId FROM
tbl_lineainvestigacion";
$responlinea = $conn->consultar($sqlconsultalinea);
$rowlinea = mysql_fetch_row($responlinea);
$produccion['nombre'] = htmlentities($produccion['nombre'], ENT_QUOTES, 'UTF-8');
//convertimos el
string a utf-8 para evitar problemas con los asentos
$id_linea = trim($rowlinea[0]);
foreach ($producciones as $produccion) {
$sqlpublicacion = "INSERT INTO tbl_publicacion(pub_intId, lin_intId, pub_strNombre,
pub_strIssn,
pub_strFecha) VALUES ('', $id_linea, '' . $produccion['nombre'] . '', '' . $produccion['issn']
. '', '' . $produccion['fecha'] . '')";
$conn->consultar($sqlpublicacion);
}
}

```

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

}
}
}
$conn->liberarConsulta();
$conn->desconectar();
echo "<br>Archivo <strong>" . $_FILES["archivo"]["name"][$i] . "</strong> indexado
correctamente";
$files = glob('files/*'); // obtiene todos los archivos
foreach ($files as $file) {
if (is_file($file)) // si se trata de un archivo
unlink($file); // lo elimina
}
} else {
echo "<br>No se ha podido mover el archivo: " . $_FILES["archivo"]["name"][$i];
}
} else {
echo "<br>No se ha podido crear la carpeta: up/" . files;
}
} else {
echo "<br>" . $_FILES["archivo"]["name"][$i] . " - NO es un archivo html";
}
}
} else {
echo "<br>No se ha subido ningun archivo";
}
}
?>
//-----grupo_investigacion.php-----//
<?php include_once 'theme/header.php'; ?>
<?php include_once 'theme/menu.php'; ?>

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<div class="container-fluid">
<div class="row">
<div class="col-sm-1 col-sm-offset-1 col-md-12 col-md-offset-0 main">
<h2 class="sub-header text-center"><i class="glyphicon glyphicon-list"></i> Listado
Grupos De
Investigaci&oacute;n</h2>
<div class="box-body">
<?php
$conn = new Conexion();
$conn->conectar();
$sql = "SELECT * FROM tbl_lineainvestigacion";
$query = $conn->consultar($sql);
?>
<table id="lista_linea_investigacion" class="table table-bordered table-hover
datatable">
<thead>
<tr>
<th>ID GRUPO</th>
<th>GRUPO DE INVESTIGACION</th>
</tr>
</thead>
<tbody>
<?php
while ($rows = mysql_fetch_array($query)) {
?>
<tr>
<td class="text-center"><?php echo $rows['lin_intId']; ?></td>
<td><?php echo $rows['lin_strNombre']; ?></td>
</tr>
<?php } ?>

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

</tbody>
</table>
</div>
</div>
</div>
</div>
<?php include_once 'theme/footer.php'; ?>
<!-- Bootstrap core JavaScript
===== -->
<!-- Placed at the end of the document so the pages load faster -->
<script
src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.3/jquery.min.js"></script>
<script>window.jQuery || document.write('<script
src="../../assets/js/vendor/jquery.min.js"></script>')</script>
<script src="bootstrap-3.3.6/dist/js/bootstrap.min.js"></script>
<!-- Just to make our placeholder images work. Don't actually copy the next line! -->
<script src="bootstrap-3.3.6/assets/js/vendor/holder.min.js"></script>
<!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
<script src="bootstrap-3.3.6/assets/js/ie10-viewport-bug-workaround.js"></script>
<script src="bootstrap-3.3.6/datatables/dataTables.bootstrap.min.js"></script>
<script src="bootstrap-3.3.6/datatables/jquery.dataTables.min.js"></script>
<script>
$(function () {
$("#lista_linea_investigacion").DataTable();
({
"paging": true,
"lengthChange": true,
"searching": true,
"ordering": true,
"info": true,

```

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

"autoWidth": true

});

});

</script>

</body>

</html>

//-----publicaciones.php-----//

<?php include_once 'theme/header.php'; ?>

<?php include_once 'theme/menu.php'; ?>

<div class="container-fluid">

<div class="row">

<div class="col-sm-1 col-sm-offset-1 col-md-12 col-md-offset-0 main">

<h2 class="sub-header text-center"><i class="glyphicon glyphicon-list"></i> **Listado
de Publicaciones**</h2>

<div class="box-body">

<div class="table-responsive">

<?php

\$conn = new Conexion();

\$conn->conectar();

\$sql = "SELECT lin_strNombre, pub_intId, pub_strNombre, pub_strIssn, pub_strFecha
FROM tbl_publicacion

INNER JOIN tbl_lineainvestigacion on tbl_publicacion.lin_intId =

tbl_lineainvestigacion.lin_intId";

\$query = \$conn->consultar(\$sql);

?>

<table id="lista_publicacion" class="table table-bordered table-hover datatable">

<thead>

<tr>

<th>GRUPO INVESTIGACION</th>

<th>ID PUBLICACION</th>

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<th>PUBLICACION</th>
<th>ISSN</th>
<th>A&Ntildeo</th>
</tr>
</thead>
<tbody>
<?php
while ($rows = mysql_fetch_array($query)) {
?>
<tr>
<td class="col-md-3"><?php echo $rows['lin_strNombre']; ?></td>
<td class="col-xs-1 text-center text-danger"><?php echo $rows['pub_intId']; ?></td>
<td class="col-md-7"><?php echo $rows['pub_strNombre']; ?></td>
<td class="text-center text-success"><?php echo $rows['pub_strIssn']; ?></td>
<td class=" text-center text-danger"><?php echo $rows['pub_strFecha']; ?></td>
</tr>
<?php
}
?>
</tbody>
</table>
</div>
</div>
</div>
</div>
</div>
</div>
<?php include_once 'theme/footer.php'; ?>
<!-- Bootstrap core JavaScript
===== -->
<!-- Placed at the end of the document so the pages load faster -->

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<script
src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.3/jquery.min.js"></script>
<script>window.jQuery || document.write('<script
src="../../assets/js/vendor/jquery.min.js"></script>')</script>
<script src="bootstrap-3.3.6/dist/js/bootstrap.min.js"></script>
<!-- Just to make our placeholder images work. Don't actually copy the next line! -->
<script src="bootstrap-3.3.6/assets/js/vendor/holder.min.js"></script>
<!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
<script src="bootstrap-3.3.6/assets/js/ie10-viewport-bug-workaround.js"></script>
<script src="bootstrap-3.3.6/datatables/dataTables.bootstrap.min.js"></script>
<script src="bootstrap-3.3.6/datatables/jquery.dataTables.min.js"></script>
<script>
$(function () {
$("#lista_publicacion").DataTable();
({
"paging": true,
"lengthChange": true,
"searching": true,
"ordering": true,
"info": true,
"autoWidth": true
});
});
</script>
</body>
</html>

//-----publicaciones_duplicadas.php-----//
<?php include_once 'theme/header.php'; ?>
<?php include_once 'theme/menu.php'; ?>
<div class="container-fluid">

```

	<p style="text-align: center;">INFORME FINAL DE TRABAJO DE GRADO</p>	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<div class="row">
<div class="col-sm-1 col-sm-offset-1 col-md-12 col-md-offset-0 main">
<h2 class="sub-header text-center"><i class="glyphicon glyphicon-list"></i> Listado
de Publicaciones
Duplicadas</h2>
<div class="box-body">
<div class="table-responsive">
<?php
$conn = new Conexion();
$conn->conectar();
$sql = "SELECT pub_intId, pub_strIssn, pub_strFecha, pub_strNombre, COUNT(*) AS
cantidad
FROM tbl_publicacion
GROUP BY pub_strNombre
HAVING cantidad > 1
ORDER BY cantidad DESC";
$query = $conn->consultar($sql);
?>
<table id="lista_publicacion" class="table table-bordered table-hover datatable">
<thead>
<tr>
<th>ID PUBLICACION</th>
<th>PUBLICACION</th>
<th>ISSN</th>
<th>A&Ntildeo</th>
</tr>
</thead>
<tbody>
<?php
while ($rows = mysql_fetch_array($query)) {

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

?>
<tr>
<td class="col-xs-1 text-center text-danger"><?php echo $rows['pub_intId']; ?></td>
<td class="col-md-8"><?php echo $rows['pub_strNombre']; ?></td>
<td class="text-center text-success"><?php echo $rows['pub_strIssn']; ?></td>
<td class=" text-center text-danger"><?php echo $rows['pub_strFecha']; ?></td>
</tr>
<?php
}
?>
</tbody>
</table>
</div>
</div>
</div>
</div>
</div>
</div>
<?php include_once 'theme/footer.php'; ?>
<!-- Bootstrap core JavaScript
===== -->
<!-- Placed at the end of the document so the pages load faster -->
<script
src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.3/jquery.min.js"></script>
<script>window.jQuery || document.write('<script
src="../../assets/js/vendor/jquery.min.js"></script>')</script>
<script src="bootstrap-3.3.6/dist/js/bootstrap.min.js"></script>
<!-- Just to make our placeholder images work. Don't actually copy the next line! -->
<script src="bootstrap-3.3.6/assets/js/vendor/holder.min.js"></script>
<!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
<script src="bootstrap-3.3.6/assets/js/ie10-viewport-bug-workaround.js"></script>

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<script src="bootstrap-3.3.6/datatables/dataTables.bootstrap.min.js"></script>
<script src="bootstrap-3.3.6/datatables/jquery.dataTables.min.js"></script>
<script>
$(function () {
$("#lista_publicacion").DataTable();
({
"paging": true,
"lengthChange": true,
"searching": true,
"ordering": true,
"info": true,
"autoWidth": true
});
});
</script>
</body>
</html>

//-----universidad.php-----//

<?php include_once 'theme/header.php'; ?>
<?php include_once 'theme/menu.php'; ?>
<div class="container-fluid">
<div class="row">
<div class="col-sm-1 col-sm-offset-1 col-md-12 col-md-offset-0 main">
<h2 class="sub-header text-center"><i class="glyphicon glyphicon-list"></i> Listado
De Universidades</h2>
<div class="box-body">
<div class="table-responsive">
<?php
$conn = new Conexion();
$conn->conectar();

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

$sql = "SELECT * FROM tbl_universidad";
$query = $conn->consultar($sql);
?>
<table id="lista_universidad" class="table table-bordered table-hover datatable">
<thead>
<tr>
<th>ID UNIVERSIDAD</th>
<th>UNIVERSIDAD</th>
</tr>
</thead>
<tbody>
<?php
while ($rows = mysql_fetch_array($query)) {
?>
<tr>
<td class="text-center text-danger"><?php echo $rows['uni_intId']; ?></td>
<td><?php echo $rows['uni_strNombre']; ?></td>
</tr>
<?php } ?>
</tbody>
</table>
</div>
</div>
</div>
</div>
<?php include_once 'theme/footer.php'; ?>
<!-- Bootstrap core JavaScript
===== -->
<!-- Placed at the end of the document so the pages load faster -->

```

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

```

<script
src="https://ajax.googleapis.com/ajax/libs/jquery/1.11.3/jquery.min.js"></script>
<script>window.jQuery || document.write('<script
src="../../assets/js/vendor/jquery.min.js"></script>')</script>
<script src="bootstrap-3.3.6/dist/js/bootstrap.min.js"></script>
<!-- Just to make our placeholder images work. Don't actually copy the next line! -->
<script src="bootstrap-3.3.6/assets/js/vendor/holder.min.js"></script>
<!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
<script src="bootstrap-3.3.6/assets/js/ie10-viewport-bug-workaround.js"></script>
<script src="bootstrap-3.3.6/datatables/dataTables.bootstrap.min.js"></script>
<script src="bootstrap-3.3.6/datatables/jquery.dataTables.min.js"></script>
<script>
$(function () {
$("#lista_universidad").DataTable();
({
"paging": true,
"lengthChange": true,
"searching": true,
"ordering": true,
"info": true,
"autoWidth": true
});
});
</script>
</body>
</html>

//-----reiniciar.php-----//
<?php
//Conexion con la base
mysql_connect("localhost", "root", "root");

```

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

//selección de la base de datos con la que vamos a trabajar

```
mysql_select_db("dbcolciencias");
```

//Creamos la sentencia SQL y la ejecutamos

```
$$SQL1 = "TRUNCATE TABLE tbl_lineainvestigacion";
```

```
$$SQL2 = "TRUNCATE TABLE tbl_publicacion";
```

```
$$SQL3 = "TRUNCATE TABLE tbl_universidad";
```

```
mysql_query($$SQL1);
```

```
mysql_query($$SQL2);
```

```
mysql_query($$SQL3);
```

```
header("location: cargar_archivo.php");
```

```
?>
```

	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

ANEXOS

Anexo 1. Manual SIA Html

1. Introducción

Este manual detalla cómo hacer uso de la aplicación web SIA-HTML (sistema de indexación de archivos html). Desarrollada bajo la modalidad de trabajo de grado.

2. Inicio

Para acceder a la aplicación es necesario disponer de un servidor web local (XAMPP, APPSERVER, WAMPSEVE) y mediante el uso de un navegador (Chrome, Mozilla, Internet Explorer). Cargar la URL: <http://localhost:3000/sia/index.php>

3. Menú

La aplicación cuenta con un el siguiente menú, el cual permite acceder a las diferentes áreas del sistema.



Figura 7. Menú de la plataforma SIA

4. Cargar Archivo

En el menú cargar archivo se accede a un directorio con archivos html proporcionados por Colciencias. Estos archivos son procesados e indexados en una base de datos.



Figura 8. Menú de carga de archivos html

El sistema da la posibilidad de procesar uno o varios archivos a la vez.

5. Universidades

En el menú universidades se tiene la opción de consultar las universidades que han sido indexadas en la base de datos.



ID UNIVERSIDAD	UNIVERSIDAD
1	Universidad De Los Andes
2	Fundación Universidad Central
3	Universidad Católica De Colombia
4	Corporación Universitaria Minuto De Dios
5	Escuela Colombiana De Ingeniería "Julio Garavito"

Figura 9. Listado de universidades

6. Grupo De Investigación

En el menú grupo de investigación se tiene la opción de consultar los grupos que han sido indexadas en la base de datos.



ID GRUPO	GRUPO DE INVESTIGACION
1	Desarrollo moral
2	Línea de Investigación en Jóvenes y Culturas Juveniles
3	Género y Cultura
4	Pavimentos y Materiales de Ingeniería
5	Grupo de Investigación en Gestión Industrial GEGI
6	Palabra, pueblo y vida
7	GIP
8	Grupo de estudios en Ingeniería de Software

Figura 10. Listado de grupos de investigación

7. Publicaciones

En el menú publicaciones se tiene la opción de consultar las publicaciones que han sido indexadas en la base de datos.



ID	PUBLICACION	ISSN	Año
1	Influencia de Programar el estudio de la psicología moral	0033-2941	2008
2	Some problems in use of the moral judgment test	0033-2941	2008
3	El computador: ¿un medio para descubrir la investigación moral?	0033-2941	2008
4	Factores predictores de bienestar subjetivo en una muestra colombiana	0033-2941	2007
5	Bienestar subjetivo en una comunidad universitaria de	0033-2941	2006
6	Modelos de la Decisión y Generalización de Reglas en una comunidad de Colombia	0033-2941	2006
7	Some problems in use of the moral judgment test	0033-2941	2006
8	Modelos de valores: Colombia	0033-2941	2004
9	A new method for understanding moral rules and moral development	0033-2941	2004
10	Modelos de valores: Ecuador	0033-2941	2004

Figura 11. Listado de publicaciones

8. Publicaciones Duplicadas

En el menú publicaciones duplicadas se tiene la opción de consultar las publicaciones que han sido indexadas más de una vez en la base de datos.



PUBLICACION	ISSN	Año
2 Some problems in use of the moral judgment test	0033-2941	2008
10 Modelos de valores: Colombia	0121-0987	2004
12 La acción moral: Contraste entre las explicaciones dadas por la filosofía y la psicología	0123-885X	2004

Figura 12. Listado de publicaciones duplicadas

9. Reiniciar DB (Base De Datos)

El botón reiniciar DB permite hacer un “delete” a los registros procesados e indexados en la base de datos. Esta función se puede utilizar indefinidamente.



Figura 13. Botón de reinicio de base de datos

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO	Código	FDE 089
		Versión	03
		Fecha	2015-01-22

FIRMA ESTUDIANTES 

Juan Fredy U.

Rafael A. Román

FIRMA ASESOR JUAN SE.

FECHA ENTREGA: _____

FIRMA COMITÉ TRABAJO DE GRADO DE LA FACULTAD _____

RECHAZADO ___ ACEPTADO ___ ACEPTADO CON MODIFICACIONES _____

ACTA NO. _____

FECHA ENTREGA: _____

FIRMA CONSEJO DE FACULTAD _____

ACTA NO. _____

FECHA ENTREGA: _____