

INFORME FINAL TRABAJO DE GRADO

Programa Ciencia de la Información y la Documentación,
Bibliotecología y Archivística

**Normalización y homologación de los árboles de información de Latinoamérica del
Norte y el Sur de la empresa PUBLICAR S.A.**

Lizeth Giovana Rodríguez Calderón

**Universidad del Quindío
Programa de Ciencias de la Información y la Documentación,
Bibliotecología y Archivística
Facultad de Ciencias Humanas y Bellas Artes
Armenia
2018**

INFORME ENTREGA FINAL

**Normalización y homologación de los árboles de información de Latinoamérica del
Norte y el Sur de la empresa PUBLICAR S.A.**

Lizeth Giovana Rodríguez Calderón

Tutor

**Angélica María Ramírez Agudelo
Ingeniera de Sistemas y Computación**

**Universidad del Quindío
Programa de Ciencias de la Información y la Documentación,
Bibliotecología y Archivística
Facultad de Ciencias Humanas y Bellas Artes
Armenia
2018**

Tabla de Contenido

INTRODUCCIÓN	7
1. Título	8
2. Área Línea y Modalidad	8
3. Área Problemática.....	9
3.1. Contexto del Problema	9
3.2. Antecedentes	10
3.3. Formulación del Problema.....	14
4. Objetivos	14
4.1. Objetivos Específicos	14
5. Justificación	15
6. Marco de Referencia	16
6.1. Marco Teórico	16
6.2. Marco Conceptual.....	18
6.3. Marco Legal.....	27
7. Metodología.....	29
7.1. Plan de Trabajo.....	29
7.2. Cronograma de Actividades	30
7.3. Procedimiento y Esquema del Trabajo de Aplicación	31
7.3.1. Acercamiento.....	31
7.3.2. Técnicas e instrumentos de recolección de información.....	34
7.3.3. Diagnóstico.....	34
7.3.4. Análisis de las Bases de Datos.	35
7.3.5. Evaluación y Seguimiento.....	35
7.3.6. Resultados y Conclusiones.....	37
7.3.7. Informe Final.....	37
7.4. Descripción del Trabajo de Aplicación	38
7.4.1. Descripción de la Empresa.....	38
7.4.2. Diagnóstico.....	38
7.4.3. Desarrollo de las Actividades.....	40

8. Recursos	45
8.1. Humanos	45
8.2. Técnicos.....	45
9. Resultados	46
10. Conclusiones.....	47
11. Recomendaciones.....	48
12. Referencias	49

Anexos

13. Anexo (Carta de Aceptación Trabajo de Aplicación)	52
14. Anexo (Consentimiento Informado)	53
15. Anexo (Carta Finalización Trabajo de Aplicación)	54
16. Anexo (Informe Resultados Trabajo de Aplicación).....	55
17. Anexo (Control Tiempo Trabajo de Aplicación).....	57
18. Anexo (Registros de Asistencia Reuniones).....	59

Listado de Tablas

Tabla 1 Plan de Trabajo	29
Tabla 2 Cronograma de Actividades	30
Tabla 3. Ejemplos iniciales de creaciones.....	31
Tabla 4. Datos finales del análisis del árbol LAN.....	40
Tabla 5. Ejemplo Base Inicial LAN	41
Tabla 6. Ejemplo de la base ajustada LAN.....	42
Tabla 7. Ejemplo del reporte de ajustes	43
Tabla 8. Ejemplo de la homologación LAS a LAN.....	44
Tabla 9. Ejemplo de secciones patrocinadas y cantidad de clientes	44

Listado de Ilustraciones

Ilustración 1. Registro fotográfico oficinas Publicar.....	32
Ilustración 2 Diferencias LAN – LAS	33
Ilustración 3 Modelo Unificado LAN - LAS	33
Ilustración 4. Flujo creación de secciones	34
Ilustración 5. Registro fotográfico presentación de avances	36
Ilustración 6. Evidencia citación a reuniones de seguimiento	36
Ilustración 7. Registro fotográfico presentación producto final	37
Ilustración 8. Esquema fase a fase	38

INTRODUCCIÓN

La normalización de información va más allá de la simple descripción de un documento o un recurso de información dentro de un sistema cualquiera, es todo el proceso que permite organizar, analizar y poner a disposición de un usuario final la información contenida en un ítem de la manera más práctica y con la capacidad de ser utilizada para generar nuevo conocimiento.

Teniendo en cuenta lo anterior este proyecto se desarrolla con la finalidad de volver usable la información comercial contenida en los árboles de información que administra PUBLICAR S.A., con el objetivo de optimizar los resultados de búsqueda realizados por los usuarios que requieren de algún producto o servicio. Además PUBLICAR S.A. es una empresa enfocada en el desarrollo de las Pyme cuya meta es aumentar la visibilidad del negocio a través de publicidad multimedia, ofreciendo soluciones digitales y gráficas para la gestión de cada empresa y entregándole oportunidades a clientes reales para mover sus negocios.

A partir de esta premisa y con la prioridad de hacer más visibles a las Pymes que pautan en el portal de las Páginas Amarillas de Publicar S.A, se planteó este trabajo de aplicación con el fin de normalizar el árbol jerárquico de temas comerciales, generando un control terminológico de las secciones en el portal, para lograr una búsqueda más óptima de los negocios por parte de los clientes potenciales de los comerciantes que pautan en este portal web.

De manera que se presenta la oportunidad de aportar y afianzar conocimientos frente al análisis y estructuración de la información, para así poder brindar a la empresa la oportunidad de mejorar la forma de entregar a sus clientes un servicio más eficiente mediante descriptores y palabras clave que permitan llegar a cada uno de ellos, aumentando así las posibilidades de sus clientes en ser encontrados mediante su publicación en medios multimedia.

1. Título

Normalización y homologación de los árboles de información de Latinoamérica del Norte y el Sur de la empresa PUBLICAR S.A.

2. Área Línea y Modalidad

El área en la cual se puede clasificar este trabajo de aplicación es el de la información dado que se busca la normalización de términos de la base de datos para la recuperación de la información y búsqueda efectiva de servicios o clientes, en cuanto a la línea de investigación este se puede clasificar en la gestión de la información, ya que con el desarrollo de este se busca ayudar a implementar estrategias de búsqueda para obtener resultados propicios para el mejoramiento de la empresa y aumentar de forma considerable la recuperación de la información mediante descriptores y palabras claves que permitan que el portal web de PUBLICAR se haga líder en el mercado multimedia atrayendo nuevos clientes y usuarios.

Lo anterior lograra obtener cambios importantes en la manera de producir, almacenar y comunicar la información de clientes aumentando la visibilidad de los negocios a través de la publicación multimedia, esto teniendo en cuenta el impacto que tiene el uso de la tecnología en estos días y que ha cambiado radicalmente los patrones culturales que afectaron la forma de vivir, trabajar, comunicar y leer de cada uno de nosotros.

La modalidad de trabajo de grado es la de aplicación que permitirá utilizar los conocimientos adquiridos en el análisis y estructuración de la información, para poder lograr aportar a la empresa PUBLICAR en alcanzar su misión de ser líder y aliado en la publicidad multimedia.

3. Área Problemática

3.1. Contexto del Problema

Publicar S.A es una empresa enfocada en el desarrollo de las Pyme cuya meta es aumentar la visibilidad del negocio a través de publicidad multimedia, ofreciendo soluciones digitales y gráficas para la gestión de cada empresa, entregándole oportunidades y clientes reales para mover sus negocios, ya que en su misión se enfoca en mover los negocios de sus clientes para que cumplan sus sueños.

Así mismo su visión apunta a convertirse en ser líderes y aliados número uno (1) en Publicidad Multimedia de las Pyme en América Latina, apostándole a ello con la práctica de sus valores como empresa, entre los cuales se encuentran *“Yo valoro a mi gente y por eso los respeto, soy proactivo, no me digan que hacer, soy innovador y nunca me conformo, lo hacemos con pasión o no lo hacemos, lograr es lo mío, cumplir metas es lo nuestro, usuarios felices traen amigos, todos los clientes todas las veces, sin excepciones, ni excusas”*.

Por otra parte cabe mencionar que la estructura orgánica de la empresa está organizada de tal manera, que satisfaga la demanda de la empresa a nivel nacional e internacional, haciendo que cada cargo cumpla la responsabilidad para llegar a los objetivos que se establecen desde la gerencia para cumplir con las expectativas de la Junta Directiva.

En este orden de ideas es importante tener en cuenta que la empresa presenta algunas falencias en la recuperación de la información de sus clientes, limitando la posibilidad de que clientes, proveedores y usuarios en general no encuentren la información rápidamente ya que el árbol de Latinoamérica del sur no se encuentra normalizado de forma adecuada, haciéndose necesario estructurarla jerárquica y semánticamente y utilizando palabras claves o términos de búsqueda que permitan encontrar de forma rápida y eficiente a clientes y servicios requeridos en la búsqueda realizada por los usuarios. Además de tener unos parámetros definidos para la creación de macrosecciones, rubros y secciones permitiendo así unificar la información y no crear registros duplicados generando inconsistencias o conflictos en el sistema.

3.2. Antecedentes

La normalización de términos es fundamental en este caso para que los clientes de Publicar puedan aumentar las posibilidades de aparecer en las búsquedas que realizan los usuarios en el portal multimedia, esto mediante la recuperación de información por términos relacionados, es por lo anterior se hace importante conocer los antecedentes de este tema para así poder entender más porque se debe hacer y cómo puede beneficiar en este caso la empresa Publicar.

Comencemos por mencionar un primer trabajo que corresponde a Palma (2008) que tiene como título “Adaptación de las directrices españolas y norteamericanas para la redacción y construcción de autoridades de materia en español para américa latina” y en el cual menciona la *Normalización de su redacción*. En principio sigue las mismas pautas que la redacción de los encabezamientos principales, tanto en lo que se refiere a la elección del singular o plural: Bibliotecas – Administración; Arquitectura – Manuscritos – Facsímiles...

En otros términos y tomando lo más importante de este trabajo para el desarrollo y objetivo de este trabajo de aplicación hay que nombrar la normalización de la redacción, que nos da la línea para poder definir las palabras adecuadas para utilizar en los árboles LAN – LAS, porque esto permitirá que los usuarios puedan obtener respuestas exactas y coincidencias de información o información relacionada con la búsqueda y es que la forma en que se escribe influye bastante en encontrar o no lo que se necesita, ya que depende del enlace y la relación de los términos para que la base de datos pueda dar la posibilidad de recuperar la información.

Un segundo trabajo es el de Martínez (2015) quien desarrollo la “construcción de un tesaurus especializado en hidrocarburos”, en la que indica que en la actualidad, la construcción de tesauros es apoyada con la utilización de aplicaciones y software especializado que automatiza las actividades que inicialmente se ejecutaban manualmente, en gran medida este cambio se debe a la migración de los tesauros de soportes físicos a ambientes electrónicos, lo cual ha facilitado su acceso y genera mayor interacción para el desarrollo y administración del tesaurus mediante el trabajo cooperativo apoyado en el uso de las tecnologías de comunicación

como la Internet.

Conviene distinguir la importancia de mencionar que este trabajo le aporta a este trabajo de aplicación una parte significativa y transcendental ya que aparecen las tecnologías de la información para la automatización de los procesos que se realizaban manualmente y permite mencionar que en este caso los árboles LAN y LAS no son otra cosa que la versión para un medio multimedia de las conocidas páginas amarillas impresas y aunque no se ha desplazado del todo el físico de estas, si se ha reducido considerablemente su impresión y distribución en la ciudad, esta situación se ha dado porque las personas ahora buscan de forma electrónica los servicios que ofrecen los clientes que pautan por medio de las páginas amarillas, dándole paso a la construcción de una estructura adecuada para la búsqueda de la información logrando la mayor visibilidad de sus clientes a los usuarios del portal web.

Un tercer trabajo es el de Pinzón (2017) quien realizó su trabajo de grado “Construcción de un tesoro especializado en comunicación social para el contexto colombiano”, en el cual señala algunas de las características de los lenguajes documentales como las siguientes:

- Ofrecen un alto grado de control terminológico, a partir de las relaciones establecidas entre los términos y de la aplicación de criterios semánticos
- Son de fácil manejo y administración, lo que nos permite una perdurabilidad del acervo terminológico empleado, con miras a una preservación y conservación del conocimiento y de su corpus terminológico
- Permiten enriquecer a partir de la inclusión de sinónimos y términos no preferentes, formas alternas de localización y recuperación de información...

Al respecto conviene decir que este trabajo de grado se relaciona con el desarrollo de este trabajo de aplicación, ya que trae a colación características importantes de la utilización de los lenguajes documentales para la normalización de términos, tales como el control terminológico que ayuda a tener claro cuáles son los términos que se deben utilizar para cada caso cuando se normaliza una base de datos, dando paso así a la relación de términos que permite la recuperación de la información de una forma ágil y rápida. Así mismo menciona

la inclusión de sinónimos y términos no preferentes que llevan a mencionar el uso de las palabras claves y el uso de descriptores para la localización y recuperación de la información de forma tal que se pueda tener la posibilidad de obtener mejores resultados en la búsqueda de la información, arrojando resultados de manera oportuna e incrementando las posibilidades de que aparezcan más coincidencias a la hora de realizarla.

Un cuarto trabajo es la tesis elaborada por Almazán (2012) que tiene como título “Herramientas para la Interoperabilidad y Normalización de datos en RI”, en uno de sus numerales que se encuentra registrado bajo el nombre de *Normalización* y la cual la define de la siguiente forma: El proceso de normalización consiste, básicamente, en la aplicación de un conjunto de atributos o dimensiones para definir adecuadamente los datos... A partir de esto se persigue: Minimizar redundancias, simplificar el mantenimiento de los datos, permitir la recuperación sencilla de los datos en respuesta a solicitudes de los usuarios, evitar datos no identificables.

De acuerdo con lo anterior este trabajo apunta directamente a la normalización en este caso de términos que se agrupan teniendo en cuenta un conjunto de atributos o información que de modo estén disponibles para los usuarios de una forma coherente, llegando a la minimización de las redundancias o términos que signifiquen lo mismo, pero se encuentren registrado o creados bajo un nombre similar y finalmente trae a colación un tema vital que es el de la recuperación de la información que corresponde a la forma en que podemos hacer que los usuarios obtengan los datos que realmente necesita.

Un quinto trabajo es el realizado por Sánchez, Paredes (2010), titulado “normalización del lenguaje documental en antropología andina boliviana, en base a un prospecto de microtesauro especializado”, hace referencia en uno de sus numerales a las palabras claves de la siguiente forma: se considera, palabra clave, al término libre extraído del texto, del título o del resumen de un documento, para ser empleado como descriptor en el proceso de la recuperación. Así mismo hace referencia a los descriptores: Coincidimos con los autores e instituciones cuando definen a un descriptor, como término normalizado o controlado, contenido en un tesauro para realizar una de las tareas más importantes del proceso documental, denominada “indización”.

He aquí en pocas palabras, cómo se podría definir las palabras claves y los descriptores que hacen parte del desarrollo de este trabajo de aplicación ya que su utilización permite que los usuarios del portal web tengan la posibilidad de encontrar información utilizando palabras asociadas al tema específico que busca, sin causar perdida o extravió de información, permitiendo de esta forma que los usuarios obtengan los resultados esperados y los clientes de Publicar puedan tener más posibilidades de ser encontrados por nuevos cliente, todo esto logrado por la utilización de una estructura adecuada de datos organizados y enlazados correctamente.

Un sexto y último trabajo para mencionar es el de Castillo (2010) que tiene como título “Nueva propuesta evolutiva para el agrupamiento de documentos en sistemas de recuperación de información” y en la cual menciona algo importante como lo es la *Reducción de palabras a su raíz*. Para reducir el número de términos distintos con los que se trabaja, el primer paso es reducir todas las palabras a su raíz (en inglés “stemming”. Por ejemplo, las palabras “medicina, “médico” y “medicinal” se reducen a la forma “medic”. Esta técnica es llamada “remoción de sufijos” (en inglés “suffix stripping”). Si bien es posible que de esta manera se lleven a la misma raíz palabras que en realidad tienen significados distintos, en general las palabras que tienen la misma raíz refieren al mismo concepto...

Finalmente con este tema se quiere reforzar el sentido que tiene el elegir bien la estructura y los términos adecuados para la normalización de una base de datos porque de esto depende que se pueda tener una buena raíz de palabras para la correcta recuperación de la información, ya que como lo menciona el autor del trabajo se debe reducir al mínimo aquellas palabras que se encuentran repetidas o que se encuentran escritas de forma diferente pero tienen el mismo significado, esta labor ayudara a que los términos raíz en la base de datos creen coherencia con las palabras o apartes específicos que permitan tanto al usuario, como al cliente y administrador hallar lo que se requiere en un menor tiempo posible.

3.3. Formulación del Problema

¿En qué forma la normalización y homologación de la información recibida o retenida por el portal web de Publicar afectaría su posicionamiento en el mercado multimedia?

4. Objetivos

Realizar la normalización y homologación de la información recibida o retenida por el portal web de publicar, estructurando jerárquica y semánticamente los términos de búsqueda o palabras clave, permitiendo una óptima recuperación de información comercial requerida por clientes, proveedores y usuarios en general de los países latinoamericanos donde opera Publicar S.A.

4.1. Objetivos Específicos

- Analizar la información recibida o retenida por el portal web de PUBLICAR S.A, para identificar los ajustes que sean necesarios para que este funcione de forma adecuada.
- Normalizar el árbol inicial LAN que cuenta con aproximadamente 9.000 términos y corresponde a Latinoamérica del Norte, agregando descriptores y palabras claves que permitan su fácil recuperación.
- Homologar los arboles de Latinoamérica del Sur – LAS, teniendo en cuenta el árbol de Latinoamérica del Norte – LAN y creando aquellas macrosecciones, rubros o secciones que se identifiquen sean necesarias y no existan.
- Estructurar el árbol LAN - LAS de acuerdo con la normalización de los términos permitiendo así que PUBLICAR se acerque a cumplir su misión de ser líder y aliado en la publicidad multimedia en América Latina.

5. Justificación

Publicar S.A es una empresa enfocada en el desarrollo de las Pyme cuya meta es aumentar la visibilidad del negocio a través de publicidad multimedia, ofreciendo soluciones digitales y gráficas para la gestión de cada empresa y entregándole oportunidades y clientes reales para mover sus negocios.

A partir de esta premisa se plante el desarrollo de un trabajo de aplicación que tiene como propósito realizar la normalización y homologación de la información recibida o retenida por el portal web de Publicar S.A., esta tarea permitirá recuperar la información de forma más rápida mediante los motores de búsqueda como Google, de tal forma que las Pyme aumentaran de forma considerable la visibilidad de su negocio a través del mercado multimedia, esto se realizara generando un control terminológico de las secciones en el portal web permitiendo obtener una estructura adecuada que enlace correctamente cada uno de los niveles del árbol de Latinoamérica del Norte y Sur.

Para poder llevar a cabo esta normalización y homologación de los árboles de Latinoamérica del Norte y Sur, se tendrá en cuenta los lenguajes controlados, la metodología para la elaboración de tesauros y la recuperación de la información, para así lograr obtener el resultado esperado, desarrollando estas dos tareas y definiendo estrategias para lograr mantener la estructura adecuada de creación de cada registro, ya que esto permitirá mantener la correcta estructura para que las Pymes puedan tener la oportunidad de ser más visibles por clientes o usuarios que visitan el portal web a diario o para buscar los servicios e información que necesitan.

Con los resultados obtenidos se tendrá la posibilidad de proponer cambios en la estructura que se venía manejando, con el fin de contar con los parámetros necesarios para mantener una base de información que permita a Publicar controlar y manejar la información de sus clientes de forma adecuada, encontrando soluciones concretas a los problemas de creación, estructura y

forma de registro de la información que se recibe y retiene, permitiendo tenerla clasificada por niveles, temas y unificando criterios.

Teniendo en cuenta lo anterior se presenta la oportunidad de aportar y afianzar conocimientos frente al análisis y estructuración de la información, para así poder brindar a la empresa la oportunidad de mejorar la forma de entregar a sus clientes un servicio más eficiente mediante descriptores y palabras clave que permitan llegar a cada uno de ellos, aumentando así las posibilidades de sus clientes en ser encontrados mediante su publicación en medios multimedia.

6. Marco de Referencia

6.1. Marco Teórico

La normalización o estandarización de términos es un proceso mediante el cual se elaboran, aplican y mejoran las normas para la creación o administración de una base de datos de información recuperada o retenida, en este caso por el portal web de la empresa Publicar, lo que lleva a mencionar a lo que Pérez (2014) hace referencia que es el trabajo realizado por la Comisión Electrotécnica Internacional – IEC, que comenzó su labor en el año de 1906 con una serie de trabajos orientados a la unificación de los métodos y a la regulación de la producción electrotécnica, por lo cual en 1938 publicaron la primera edición del Vocabulaire électro-technique international.

En relación con este tema Pérez (2014) trae a colación la aparición del Organismo Internacional de Normalización (ISO), al interior del cual se creó un comité técnico para la normalización terminológica que permitió la publicación de las primeras normas sobre terminología en 1968, relacionado con el ISO se pueden mencionar algunas de las normas que atañan a la producción, la armonización y la gestión terminológica.

De acuerdo con lo anterior es importante mencionar que la normalización de términos a pesar de no ser un tema nuevo, no se ha visto tan impulsado y es por eso que con la

información mencionada anteriormente es de anotar que la normalización de términos es una labor que está respaldada en normas que permiten llevar a cabo de forma adecuada este proceso obteniendo resultados satisfactorios en la recuperación de la información, logrando así que las personas puedan encontrar la información que requieren de forma inmediata.

Pérez (2014) indica que la creación de estas normas ha permitido ayudar a organizaciones y a empresas a crear, homogeneizar y gestionar sus terminologías, así como a mejorar la información contenida en los estándares técnicos propuestos. Todo esto con el fin de brindar al usuario la oportunidad de obtener información en tiempo real y recuperarla mediante palabras claves que atan toda la estructura de una base de datos permitiendo con solo ingresar un término conocido tener la mayor cantidad de información asociada a este.

Es así como Wüster (2010) en su libro *Introducción a la Teoría General de la Terminología* y a la *Lexicografía Terminológica* señala que se pueden establecer dos tipos de relación entre los conceptos las relaciones lógicas y las relaciones ontológicas, la primera se da cuando existe una relación directa entre los términos y la segunda son las relaciones indirectas entre los términos. Es por lo anterior que es importante mencionar que Cabré (1993) enfatiza en la importante labor que realiza la terminología como vehículo de comunicación entre especialistas y profesionales, para lograr el acceso a la misma sin dejar de lado las necesidades de los usuarios.

Es oportuno ahora definir los lenguajes documentales que Gavilán (2009) en su publicación *“Lenguajes documentales Principales tipos de clasificación Encabezamientos de materia, descriptores y tesauros”* los establece como la comunicación entre el documento original y el usuario, haciendo de puente o intermediario entre las informaciones contenidas en los documentos y la información solicitada por los usuarios. Dando paso a un poco de la historia de estos y es que aunque no se crea los lenguajes documentales existen desde el siglo XIX, época en la cual intentaban recopilar de manera sistemática el contenido de que tratan los libros y así ir creando una clasificación del saber humano.

De la misma manera en el siglo XX comenzaron a surgir los lenguajes especializados los cuales permitirían agrupar palabras por temática e ir estructurando así una forma clara para recuperar la información necesaria sin tener que buscar en varias fuentes de información o tardando varias horas o días intentando obtener resultados satisfactorios en la búsqueda.

De esta manera es conveniente mencionar que el ICONTEC es una entidad que se encuentra creada desde 1963 y que es un Organismo Asesor y Coordinador en Normalización, por lo tanto genera normas técnicas para la estandarización de procesos mediante la formulación, publicación e implementación de normas en Colombia, además de representarla ante los organismos de normalización internacionales y regionales como la ISO (International Organization for Standardization); IEC (International Electrotécnica Commission) COPANT (Comisión Panamericana de Normas Técnicas).

En resumidas cuentas puede notarse que la normalización de términos es una labor con la cual se logra obtener bases de datos más depuradas, estandarizando palabras claves y empleando la terminología correcta para la recuperación de la información, igualmente es de resaltar que la normalización de términos es una labor que se viene desarrollando desde hace muchos años, pero al igual que otras labores no son tan conocidas, pero son realmente beneficiosas para el funcionamiento de las empresas y en el caso de este trabajo de aplicación atraer más clientes, usuarios y proveedores al portal de Publicar.

6.2. Marco Conceptual

La normalización tiene como objetivo lograr aplicar y mejorar las normas que se utilizan para distintas actividades científicas, industriales o económicas, con el fin de ordenarlas y mejorarlas, esto nos lleva a mencionar la terminología la cual desempeña un papel decisivo en la comunicación técnica y científica y que permite la transferencia de conocimientos, entendiendo el lenguaje con el cual se lleva a cabo esta transferencia. Así mismo teniendo en cuenta lo anterior es importante dar a conocer conceptos importantes y relacionados para el desarrollo de este trabajo de aplicación entre los cuales encontramos los siguientes:

Rubio (2004) menciona que el **Análisis de Contenido**. La información implica la identificación y representación del contenido de los textos mediante dos técnicas fundamentales: la indización y el resumen.

Rubio (2004) menciona que el **Análisis Documental**. Es un trabajo mediante el cual por un proceso intelectual extraemos unas nociones del documento para representarlo y facilitar el acceso a los originales. Analizar, por tanto, es derivar de un documento el conjunto de palabras y símbolos que le sirvan de representación.

Según Senn (1992) el **Análisis Estructurado**. Es un método utilizado para el análisis de sistemas automatizados o manuales, que conduce al entendimiento del sistema actual, y así proponer el desarrollo de especificaciones de sistemas nuevos o para efectuar modificaciones a los ya existentes.

Henao (2014) menciona que el **Análisis Gramatical**. Lo componen tres tareas, estas son: **Primera** consiste en clasificar las palabras conforme a sus funciones más generales. Es fácil ver que algunas palabras tienen igual fuerza y capacidad para combinarse en enunciados inteligibles, **Segunda** consiste en describir los modos más o menos sistemáticos en que las palabras pueden modificarse agregando los morfemas que les son propios consecuentes cambios en el significado y **Tercero** consiste en enunciar la sintaxis, es decir establecer reglas de combinaciones admisibles o gramaticalmente correctas de los morfemas. Parte de esta tarea consiste en reemplazar el criterio de las pausas naturales por los criterios formales para el análisis de las oraciones.

Torres (2016) menciona que la **Base de Datos**. Es una colección de información útil organizada en una manera específica. Por ejemplo para llevar el control de los teléfonos de personas podría ser útil una base de datos que constaría de los datos de la tabla agenda.

Torres (2016) menciona que en las **Bases de Datos de Redes**. La redundancia se reduce al máximo, no se pierde información debido a que una parte no se suministre o a que un proveedor no suministre alguna parte.

Torres (2016) menciona que las **Bases de Datos Jerárquicas**. Permiten “ver” una base de datos como compuesta de un conjunto de relaciones padre – hijo.

Torres (2016) menciona que las **Bases de Datos Relacionales**. “Ve” a los datos como un conjunto de tablas (Archivos) constando de un conjunto de renglones (Registros) y cada renglón constando de un conjunto de columnas (Campos). La relación entre las tablas (Archivos) se establece solo por nombres de atributos comunes.

Gavilán (2009) menciona que las **Clasificaciones Especializadas**. Profundizan en campos específicos del conocimiento, por lo que los términos utilizados corresponden a un área temática en particular. Algunas de las más conocidas clasificaciones de este tipo son la de la National Library of Medicine (EEUU), la Clasificación Decimal Astronáutica y la Clasificación de la OCDE.

Parra (2015) menciona que las **Clasificaciones Facetadas**. Es un tipo de clasificación se usa para organizar conjuntos de objetos que pueden estar en un sitio web, productos, documentos, etc., que son lo suficientemente homogéneos como para ser descritos por sus atributos o propiedades (facetas y categorías) y sus valores (pertenencia a categorías), en otras palabras, por un conjunto de metadatos facetados.

Gavilán (2009) menciona que las **Clasificaciones Monojerárquicas o Jerárquicas**. También denominadas analítico-enumerativas. Sus clases, y subclases se organizan en forma arborescente, dando lugar una serie de divisiones y subdivisiones dependientes unas de otras. Los sistemas de estructura jerárquica pueden ser especializados, cuando sólo se ocupan de una rama del saber o enciclopédicos, cuando abarcan el universo de los conocimientos.

Gavilán (2009) menciona que los **Descriptor**s. Son términos del lenguaje natural que por medio de un proceso de selección llega a formar parte del vocabulario de un tesauro. La condición que debe superar un término para adquirir la condición de descriptor es que sea el más representativo del concepto que se quiere representar dentro del tesauro.

Castro (2016) menciona que los **Diagramas de Contexto**. Van desde lo general hasta lo específico. Este diagrama debe ser un resumen, debe incluir las entradas básicas, el sistema general y las salidas. Así convertirse en la conceptualización del sistema de manera general.

Castro (2016) menciona que el **Diccionario de Datos**. Es una lista de todos los elementos que forman parte del flujo de datos en todo el sistema. Los elementos más importantes son flujos de datos, almacenes de datos y procesos.

Lancaster (2002) menciona que la **Especificidad del Vocabulario**. Controla directamente el nivel de precisión que pueda conseguirse en la búsqueda porque controla el tamaño de las categorías a las que pertenecen los documentos.

Castro (2016) menciona que según Kendall (2002) el **Flujo de Datos**. Facilita la determinación de requerimientos La intención de analizar un sistema es entender los requerimiento de información de los usuarios, por lo tanto se debe conceptualizar cómo fluyen los datos a través de la organización, los procesos o la transformación que los datos puedan llegar a sufrir.

La Escuela Pública Digital (s.f.) menciona que el **Gestor de Base de Datos o Sistema de Gestión de Base de Datos (SGBD o DBMS)**. Es un software que permite introducir, organizar y recuperar la información de las bases de datos; en definitiva, administrarlas.

Henao (2014) menciona que la **Gramática Descriptiva**. Se encarga de analizar el uso de una lengua y el uso de las palabras que una comunidad da a ella.

Henao (2014) menciona que la **Gramáticas Formales**. Se cuenta una de las más innovadoras como la lingüística computacional en la que se plantea una sintaxis del lenguaje de programación.

Henao (2014) menciona que la **Gramática Funcional**. Estudia cómo se emplean las palabras, los sintagmas y las oraciones y la adecuación de estos elementos según sea el contexto y la situación comunicativa en que se utilicen.

Rodríguez (2018) **Homologación**. Agrupación de términos, ya sea por sus características, especificaciones o significado.

Rubio (s.f.) menciona que **Indizar**. Es extraer una serie de conceptos que responden a los temas tratados en el documento, y que servirán como puntos de acceso para su recuperación.

Henao (2014) menciona que la **Intertextualidad**. Es el conjunto de relaciones que acercan un texto determinado a otros textos de variada procedencia, o en sus propias palabras "todo texto es la absorción o transformación de otro texto".

Gavilán (2009) menciona que los **Lenguajes Controlados**. Son lenguajes controlados los lenguajes documentales propiamente dichos: tesauros, listas de encabezamientos de materia y clasificaciones. Presentan un vocabulario previamente elaborado, y admiten un limitado número de modificaciones en el momento de su utilización. El vocabulario controlado proporciona al usuario un punto de búsqueda, en vez de dos o más, y reduce la posibilidad de que la búsqueda sea incompleta. Sin embargo, puede perderse alguna información debido a la falta de especificidad y errores en el análisis.

Parra (2015) menciona que el **Lenguaje Documental**. Sistema artificial de signos utilizados en las operaciones de indización que permite la representación de contenido documental para su posterior recuperación, sirviendo como medio para la interrogación, recuperación y difusión de información pertinente para el usuario.

Parra (2015) menciona que el **Lenguaje de Análisis y Recuperación de Información – LARI**. Permiten la recuperación de la información, previamente almacenada, por medio de la realización de una serie de consultas a los documentos contenidos en la base de datos.

Gavilán (2009) menciona que los **Lenguajes de Estructura Asociativa o Combinatoria**. Son aquellos que permiten la libre combinación de términos entre sí, eliminando la rigidez de los lenguajes de estructura jerárquica. A la estructura combinatoria o asociativa responden, por ejemplo, las listas de encabezamientos de materia y los tesauros, si bien es cierto que en la estructura de estos últimos participa también la estructura jerárquica. Estos lenguajes tienen una presentación alfabética y los conceptos se asocian o combinan libremente entre sí para expresar el tema o los temas.

Gavilán (2009) menciona que los **Lenguajes de Estructura Jerárquica**. A la estructura jerárquica, también denominada arbórea o sistemática, responden determinadas clasificaciones (las clasificaciones jerárquicas): por ejemplo, la Dewey y la CDU. En ellas los conceptos se distribuyen en clases o categorías, donde dependen unos de otros. Estos lenguajes van de lo general a lo particular. Útiles en fondos enciclopédicos, resultan poco operativos para recuperar información, son sistemas muy rígidos y difíciles de actualizar, pueden servir para organizar colecciones.

Gavilán (2009) menciona que los **Lenguajes de Estructura Sintáctica**. Surgen por la cantidad de ruido que producen los lenguajes combinatorios, que pueden poseer una red de relaciones semánticas pero no tienen gramática. La estructura sintáctica permite concretar y ordenar los conceptos. Este tipo de lenguaje requiere un cierto grado de automatización en los procesos tanto de indización como de recuperación, por lo que su aparición está asociada al desarrollo tecnológico. Según su complejidad, se pueden diferenciar dos tipos: lenguajes sintácticos de gramática simple y lenguajes sintácticos de gramática elaborada. De momento, los que aparecieron a finales de los años sesenta, como Syntol (Syntagmatic Organization Language) no han tenido demasiado éxito.

Gavilán (2009) menciona que los **Lenguajes libres**. Son lenguajes fundados en el principio de postcoordinación, se componen de términos no predefinidos que se va generando a partir de la realización de procesos de indización. Lenguajes altamente pertinentes, coherentes, de bajo coste y que proporcionan una enorme cantidad de resultados cuando se realizan búsquedas. De este tipo son las listas de descriptores libres y las listas de palabras clave

(Uniterm, Índices permutados, Precis ...). Los lenguajes libres no son propiamente lenguajes documentales puesto que para que reciban este nombre el vocabulario ha de estar controlado.

Gavilán (2009) menciona que los **Lenguajes Precoordinados**. Combinan los términos en el momento de la descripción. Son lenguajes precoordinados las clasificaciones y las listas de encabezamientos de materia. En estos sistemas las distintas nociones o conceptos que se unen para expresar una materia o un tema ocupan un lugar determinado, es decir se introducen en el momento de la indización en un orden previamente establecido y la recuperación habrá de hacerse secuencialmente, siguiendo ese orden. Suelen ser muy precisos, pero también muy rígidos.

Gavilán (2009) menciona que los **Lenguajes Postcoordinados**. Combinan los términos en el momento de la recuperación. Son lenguajes postcoordinados los tesauros (lenguaje utilizado para la indización por descriptores), y si incluyéramos los lenguajes libres podemos mencionar también las listas de descriptores libres y las listas de palabras clave. En estos sistemas las nociones o conceptos que se extraen en la indización para expresar el tema o los temas del documento tienen todas la misma categoría y no se expresarán en ningún orden determinado. Serán los intereses de los usuarios en la fase de recuperación los que obliguen a la combinación de conceptos que se realizará utilizando los operadores booleanos, los operadores de comparación, etc.

Gavilán (2009) menciona que las **Palabras Clave o los Unitérminos**. Se diferencian de los descriptores en que, a pesar de compartir la procedencia del lenguaje natural y de tener la función de representar conceptos, las palabras claves carecen del control estricto de la sinonimia, de las relaciones asociativas y, mucho menos de las jerarquías.

Gavilán (2009) menciona que las **Presentación Sistemática de los Tesauros**. Contiene los descriptores estructurados en función de categorías. Los descriptores se sitúan en sus áreas contextuales donde son jerarquizados por categoría semántica y dotados de distintos tipos de relaciones. El establecimiento de categorías se puede realizar a partir de facetas o de campos temáticos.

Molina (2011) menciona que la **Recuperación de Información**. Es proporcionar información relevante al usuario para satisfacer una necesidad de información. Este proceso comienza cuando una persona necesita información sobre un tema especial, ya sea que desee buscarla por sí misma o necesite la ayuda de un referencista, a esto es lo que llamamos “solicitud de búsqueda de información. Para esto se requiere de un Sistema de Recuperación de Información.

Gavilán (2009) menciona que las **Relaciones Asociativas entre Descriptores**. Es el tipo de relación más importante del tesauro por ser la que le define como lenguaje de estructura combinatoria, hecho que le diferencia de otros. Es una relación recíproca que identifica descriptores afines, que no jerárquicos entre sí, sean o no de una misma familia, en una dirección horizontal.

Gavilán (2009) menciona que las **Relaciones Definitivas o Aclaratorias entre Descriptores**. Se utilizan para delimitar la aplicación de un descriptor. Puede describir el significado por el cual se va a usar un término con sinonimia, acotar un significado a una única opción excluyendo algún sentido, eliminar ambigüedades, desarrollar siglas o abreviaturas, etc. Se utiliza con las notaciones NA (Nota de Alcance), en inglés SN (Scope Note). Pueden ser de varios tipos: definición completa; definición restringida; inclusión de términos en relación jerárquica con el descriptor; descripción de cómo debe usarse el término en la indización; e historia del descriptor.

Gavilán (2009) menciona que las **Relaciones de Equivalencia entre Descriptores**. También llamadas relaciones de sustitución o preferenciales. Dirigen desde el sinónimo, cuasi sinónimo o antónimo donde se centra la demanda, hacia el descriptor autorizado o término preferencial. Esta relación supone una traducción del término natural (libre) al artificial (normalizado). Las notaciones de esa relación son: USE y UP (Usado por) o en inglés USE y UF (Use For).

Gavilán (2009) menciona que las **Relaciones Jerárquicas entre Descriptores**. Define los descriptores más genéricos y más específicos, con sus correspondientes clases, especies,

familias, partes, etc. A deferencia de los sistemas jerárquicos tradicionales, esta relación se establece en función del resto de las relaciones de un concepto, sin las cuales la jerarquía no siempre sería posible: un descriptor puede tener términos más genéricos o más específicos relacionados con él porque a su vez cuenta con relaciones asociativas, de sinonimia o definitorias que le permiten tener cualquiera de las tres relaciones restantes, sin que al mismo tiempo, ninguna sea excluyente de otras.

Castro (2016) menciona que la **Salida de Datos**. Es aquella que se entrega al usuario por medio de un Sistema de Información, donde los datos requieren un procesamiento extenso antes que se conviertan en salida adecuada.

Henao (2014) menciona que la **Semiótica**. Es la jerarquía cualquiera de cuyos componentes puede sufrir un ulterior análisis en clases definidas por una relación recíproca, de modo que cada una de estas clases puede sufrir un análisis en derivados definidos por una mutación recíproca

Henao (2014) menciona que la **Semiótica de la Comunicación**. Es aquella que considera a la realidad signficante como un proceso de Comunicación ente un Emisor y un Receptor.

Henao (2014) menciona que los **Sistemas de Comunicación**. Pueden diferir respecto de los medios, los mensajes expresados o expresables y respecto de las convenciones para asociar las señales con los significados.

Cruz (s.f.) menciona que los **Sistemas Manejadores de Bases de Datos – SMBD**. Son conocidos también como bases de datos electrónicas, se usan ampliamente para organizar y manipular grandes volúmenes de datos de las empresas.

Gavilán (2009) menciona que el **Tesaurus**. Es un prototipo de lenguaje de indización y recuperación controlado, que se basa en la postcoordinación de sus descriptores. Además representa de manera unívoca los conceptos de los documentos evitando así los problemas de homonimia, sinonimia, polisemia (relaciones de significado), así como el establecimiento de

las relaciones jerárquicas y de relación entre los descriptores que lo componen.

Parra A. (2015) menciona que los **Tesauros Arborescentes**. Estructurados en forma ramificada de tal manera que cada descriptor puede tener una cabeza con varias dependientes, relacionándose siempre por vía ascendente o descendente.

Parra A. (2015) menciona que los **Tesauros Automatizados**. Se incorporan a una base de datos para que el usuario pueda conocer cuál es el contenido del vocabulario a utilizar, su redacción y aprobación.

Vilares, J. Barcala, M. Alonso, M. (s.f.). mencionan que las **Variantes Sintácticas**. Son producto de la reflexión de palabras individuales y de la modificación de la estructura sintáctica del término original.

Vilares, J. Barcala, M. Alonso, M. (s.f.). mencionan que las **Variantes Morfosintácticas**. Al menos una de las palabras con contenido del término original transforma en otra que deriva de la misma raíz morfológica.

6.3. Marco Legal

Aunque la normalización de términos no es un tema nuevo, si es un tanto desconocido para muchas personas o entidades que probablemente no lo utilizan porque no saben cuáles son sus beneficios y mucho menos llegan a imaginar que se encuentra reglamentado. Con relación a lo anterior se puede mencionar que se han creado normas y organizaciones, las cuales han logrado brindar información, normar este tema y orientar a los profesionales de Ciencias de la información que lo implementan y usan en sus actividades diarias.

Teniendo en cuenta lo anterior es importante mencionar algunas de las normas que se han creado para normalizar la información y obtener bases de datos con los parámetros suficientes para optimizar tiempos de búsqueda por clientes y usuarios, a continuación se relaciona:

Por una parte encontramos las normas creadas por la Organización Internacional de Normalización – ISO, las cuales son:

- ✓ **ISO 639: 1988** - Código para la representación de nombres de idiomas
- ✓ **ISO 704: 1987** - Principios y métodos de terminología
- ✓ **ISO DIS 860: 1993** - Armonización internacional de conceptos y términos
- ✓ **ISO 1087: 1990** - Vocabulario de terminología
- ✓ **ISO 2145: 1978** - Documentación - Numeración de divisiones y subdivisiones en documentos escritos
- ✓ **ISO 10241-1:2011** - Entradas terminológicas en estándares.
- ✓ **ISO 2788:1986** - Directrices para el establecimiento y desarrollo de tesauros monolingües.
- ✓ **ISO 5964:1985** - Principios rectores para la elaboración y el desarrollo de tesauros multilingües
- ✓ **ISO 15188: 2001** - Directrices para el desarrollo de terminología normalizada.

Así mismo el Instituto Colombiano de Normas Técnicas y Certificación – ICONTEC, en Colombia, publicó en 1980 la **NTC 1544 - Documentación. Indización. Terminología.**, la cual tiene como objetivo precisar y definir los términos utilizados en indización, con esta se apoya la normalización de los términos para las bases de datos brindando la información necesaria para su desarrollo de la normalización de bases de datos y lograr así reducir la duplicidad de información y mejorar la productividad de la empresa o del sistema según sea el caso.

También se puede encontrar la **NTC 1476:1998** – Guía para el establecimiento y desarrollo de tesauros monolingües, en la cual se menciona que el índice temático como medio para identificar y recuperar documentos depende de un lenguaje de indización bien construido. Esto se aplica a todo sistema donde se deben tomar decisiones intelectuales para la selección de términos de indización, incluyendo aquellos sistemas que utilizan un computador para almacenar y manipular términos o para identificar documentos asociados con términos o con combinaciones de términos asignados por un indizador.

Finalmente se puede mencionar que aún falta documentar un poco más el tema y crear más normas pero con las existentes se puede realizar un buen trabajo de normalización de información para poder brindar a los usuarios una mejor estructura basada en los parámetros dados por la normatividad vigente, permitiendo así obtener la información rápidamente y de forma organizada, pudiendo brindar a las empresas o personas la oportunidad de visualizar lo que está buscando. Tanto así que han servido como insumo guía para el desarrollo de este trabajo de aplicación para entregar un producto adecuado para lograr el objetivo del mismo.

7. Metodología

7.1. Plan de Trabajo

Para el desarrollo de este trabajo de aplicación se definieron unos objetivos específicos y para dar cumplimiento a estos se plantearon las acciones, recursos y evidencias para poder obtener los resultados esperados con su desarrollo. Lo anterior se puede evidenciar en la siguiente tabla:

Tabla 1 Plan de Trabajo

PLAN DE TRABAJO			
OBJETIVOS ESPECÍFICOS	ACCIÓN	RECURSOS	EVIDENCIAS
Analizar la información recibida o retenida por el portal web de Publicar S.A, para identificar los ajustes que sean necesarios para que este funcione de forma adecuada.	Análisis del árbol LAN por nivel (I, II y III)	Analista Computador Base de datos	Reportes de ajustes identificados por nivel (I, II y III) Registros de Asistencia Reuniones (Ver anexo 18 - Registros de Asistencia Reuniones)
Normalizar el árbol inicial LAN que cuenta con aproximadamente 13.000 términos y corresponde a Latinoamérica del Norte, agregando descriptores y palabras claves que permitan su fácil recuperación.	Ajustar el árbol LAN por nivel (I, II y III) teniendo en cuenta el análisis realizado	Analista Computador Base de datos	Árbol LAN con los ajustes detectados en el análisis

PLAN DE TRABAJO			
OBJETIVOS ESPECÍFICOS	ACCIÓN	RECURSOS	EVIDENCIAS
			Registros de Asistencia Reuniones (Ver anexo 18 - Registros de Asistencia Reuniones)
Homologar los arboles de Latinoamérica del Sur – LAS, teniendo en cuenta el árbol de Latinoamérica del Norte – LAN y creando aquellas macrosecciones, rubros o secciones que se identifiquen sean necesarias y no existan.	Homologar los arboles de Latinoamérica del sur – LAS, teniendo en cuenta el árbol ajustado de Latinoamérica del norte – LAN	Analista Computador Base de datos	Árboles LAS homologado al LAN Registros de Asistencia Reuniones (Ver anexo 18 - Registros de Asistencia Reuniones)
Estructurar el árbol LAN - LAS de acuerdo con la normalización de los términos permitiendo así que Publicar se acerque a cumplir su misión de ser líder y aliado en la publicidad multimedia en América Latina.	Ajustar el árbol LAN – LAS teniendo en cuenta la homologación de términos realizada	Analista Computador Base de datos	Árbol LAN – LAS Unificado Registros de Asistencia Reuniones (Ver anexo 18 - Registros de Asistencia Reuniones)

7.2. Cronograma de Actividades

Para la realización de este trabajo de aplicación se tendrá en cuenta un periodo comprendido entre el 1 de marzo al 27 de abril del 2018 y una dedicación de cuatro (4) horas diarias, para un total de 9 semanas y 136 horas para su desarrollo, tal como se puede observar en la siguiente tabla:

Tabla 2 Cronograma de Actividades

CRONOGRAMA									
ACCIONES	MARZO (4 horas diarias)					ABRIL (4 horas diarias)			
	1-2	5-9	12-16	19-23	26-30	2-6	9-13	16-20	23-27
1. Análisis del árbol LAN por nivel (I, II y III)	X	X	X						
2. Ajustar el árbol LAN por nivel (I, II y III) teniendo				X	X				

CRONOGRAMA									
ACCIONES	MARZO (4 horas diarias)					ABRIL (4 horas diarias)			
	1-2	5-9	12-16	19-23	26-30	2-6	9-13	16-20	23-27
en cuenta el análisis realizado									
3. Homologar los arboles de Latinoamérica del Sur – LAS, teniendo en cuenta el árbol ajustado de Latinoamérica del Norte – LAN					X	X	X		
4. Ajustar el árbol LAN – LAS teniendo en cuenta la homologación de términos realizada								X	X
Total de tiempo de dedicación para el desarrollo del TA 136 horas									

7.3. Procedimiento y Esquema del Trabajo de Aplicación

Para el desarrollo de este trabajo de aplicación se definió un procedimiento el cual será descrito a continuación con el fin de detallar cada uno de sus pasos y se pueda entender de que forma se llevaron a cabo cada una de las actividades:

7.3.1. Acercamiento.

En este paso se realizó un acercamiento con las personas encargadas (*Ver anexo 18 - Registros de Asistencia Reuniones*) de la base de datos de la información recibida o retenida por el portal web de Publicar trabajando eventualmente en sus oficinas, avanzando así en la identificación de los problemas que a simple vista eran notorios como la duplicidad de términos creadas con diferentes ID o algunos con el mismo ID, así como se puede notar con los siguientes ejemplos:

Tabla 3. Ejemplos iniciales de creaciones

ID	Sección	ID	Sección	ID	Sección
5281	Tapas de Plastico	1377	Cocina Gourmet	4531	Recuperación de Activos
5281	Tapas Plasticas	1377	Comida Gourmet	32924	Activos Recuperación de



Ilustración 1. Registro fotográfico oficinas Publicar.

Por otra parte las personas encargadas nos suministran la información necesaria para comprender la estructura de la base de datos, estos datos fueron los siguientes:

- ✓ **Definición macrosección.** Es el campo más general, hace referencia a sectores económicos, pero no a una actividad comercial específica.
- ✓ **Definición Rubro.** Hace mención a una actividad comercial, segmenta los sectores económicos en campos más restringidos, lo que genera mejores resultados para el usuario en relación al nivel de Macrosección.
- ✓ **Definición Sección.** Se refiere a una actividad económica más específica; es el más especializado de los tres niveles del Árbol de Secciones.
- ✓ Estructura de los árboles de Latinoamérica del Norte y Sur:

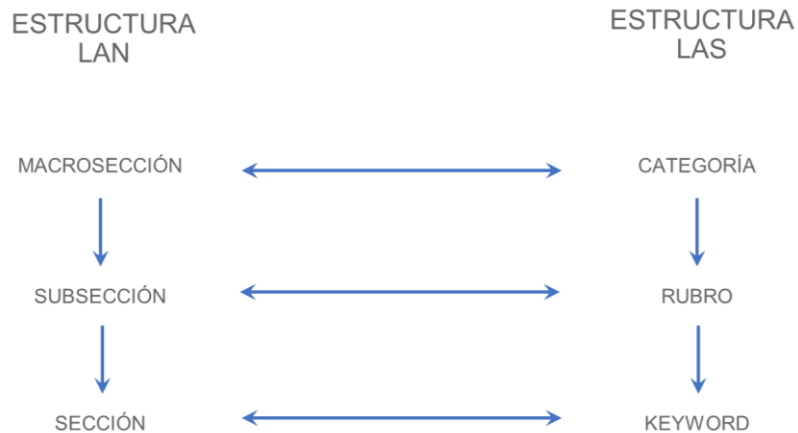


Ilustración 2 Diferencias LAN – LAS

- ✓ Modelo unificado para la estructura del árbol final LAN – LAS:

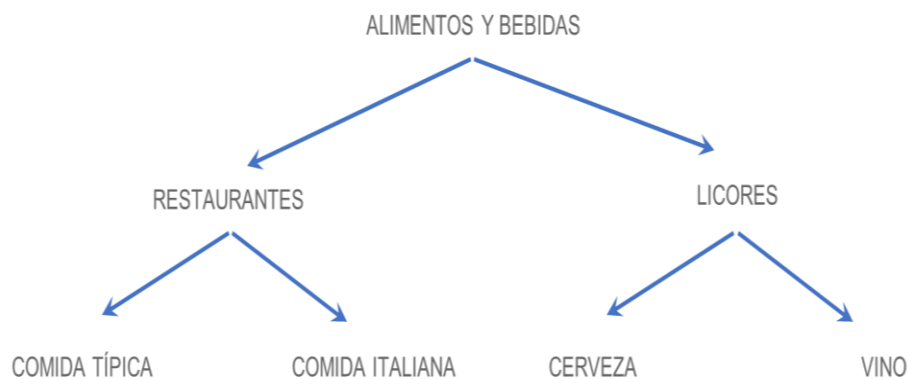


Ilustración 3 Modelo Unificado LAN - LAS

✓ Flujo de creación de secciones:

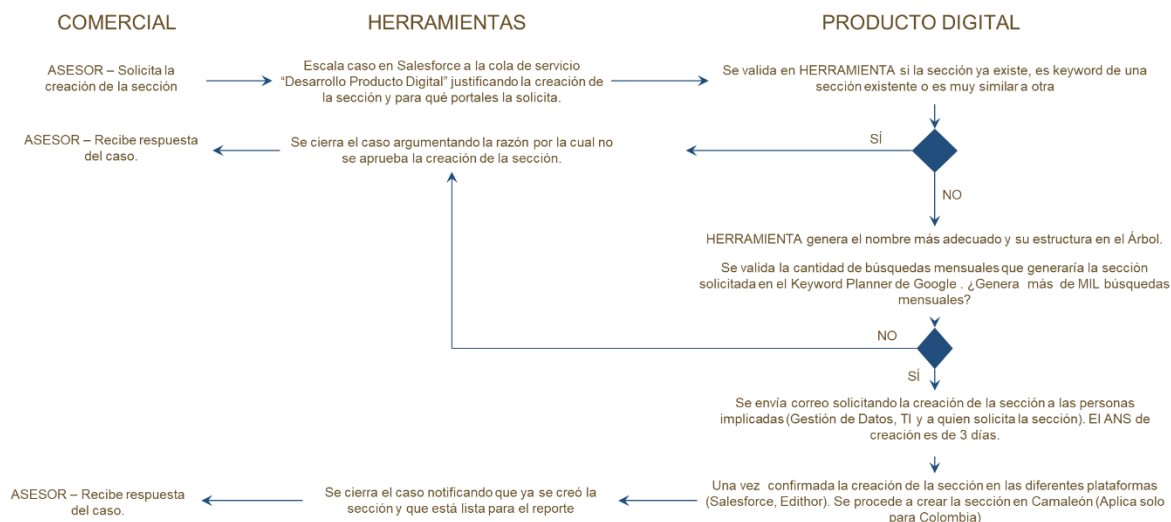


Ilustración 4. Flujo creación de secciones

7.3.2. Técnicas e instrumentos de recolección de información.

En este paso la técnica para la recolección de la información es la observación que es un método que permite la percepción directa del objeto de investigación en este caso la estructura del árbol LAN y LAS para posteriormente tomar las decisiones respectivas frente a la información contenida en estos, dando paso al análisis de la información y permitiendo normalizar y homologar los términos que haya lugar en su momento. Además de brindar la posibilidad de tener una idea general del estado de la base de datos y dando paso a la realización de los ajustes necesarios dejando registro de estos mismos mediante un reporte que posteriormente será utilizado para realizar los cambios por el área de tecnología de la empresa.

7.3.3. Diagnóstico.

Después de aplicar la técnica de recolección de información, se puede obtener una clara visión del estado de los árboles LAN - LAS e identificar las falencias que estos tienen que no permiten que la información contenida en el portal web, pueda ser recuperada de forma fácil en los motores de búsqueda web, dando paso así a que las Pymes puedan tener una mejor

visibilidad para sus clientes y así mejorar la posibilidad de que tengan una base de datos que cumpla con los estándares y necesidades requeridas.

7.3.4. Análisis de las Bases de Datos.

Después de realizar el diagnóstico de la base de datos de los árboles LAN – LAS, se procede a realizar el análisis de la misma para identificar los problemas que tiene y ajustarla de una forma adecuada para lograr tener un producto que cumpla con los parámetros dados por las normas para la normalización de términos y así poder generar el reporte de los ajustes necesarios a realizar en la base de datos del portal web de Publicar.

7.3.5. Evaluación y Seguimiento.

Se llevaron a cabo sesiones en las cuales se presentaban los avances (*Ver anexo 18 - Registros de Asistencia Reuniones*) del análisis de la base de datos y la estructuración de la misma, dando paso así al feedback o retroalimentación de los casos identificados durante el análisis para poder obtener un producto final adecuado con las necesidades de la empresa Publicar y apuntando al objetivo de este trabajo de aplicación.

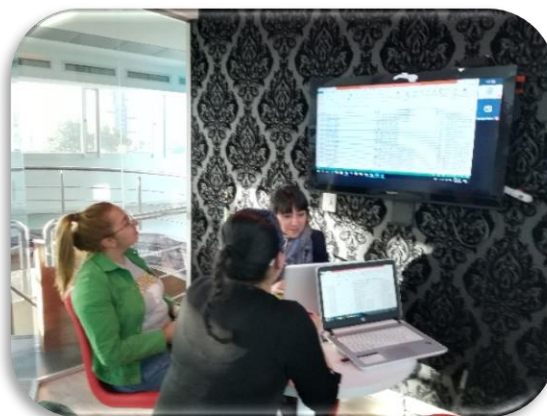
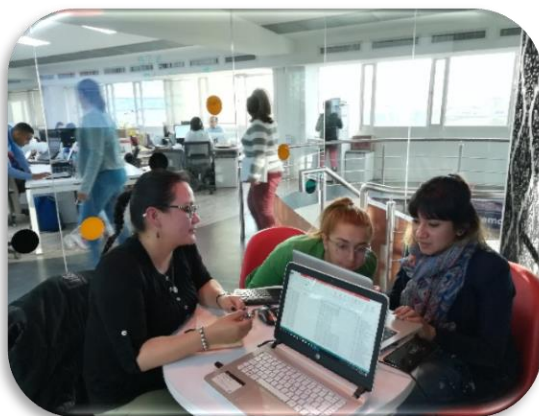




Ilustración 5. Registro fotográfico presentación de avances

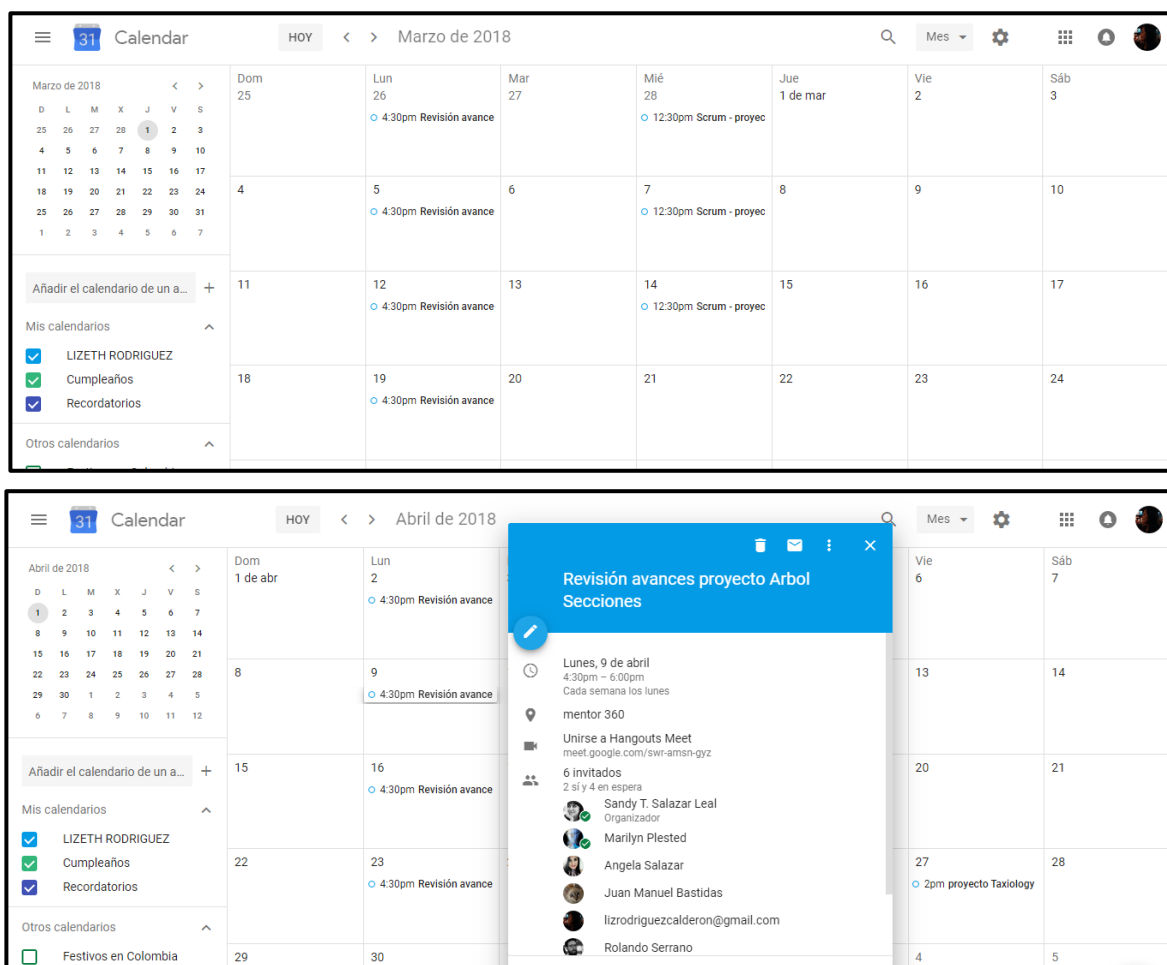


Ilustración 6. Evidencia citación a reuniones de seguimiento

7.3.6. Resultados y Conclusiones.

Una vez realizado el análisis de los árboles LAN – LAS, se dio paso a la muestra de los resultados de la normalización y homologación de estos, para así presentar a las personas encargadas el producto final (*Ver anexo 18 -Registros de Asistencia Reuniones*) y las conclusiones que se obtuvieron del desarrollo de este trabajo de aplicación para la empresa.



Ilustración 7. Registro fotográfico presentación producto final

7.3.7. Informe Final.

Se realizó un informe final en el cual se presentó el desarrollo del trabajo de aplicación y se evaluó por ambas partes el desempeño e impacto que tuvo para la empresa.



Ilustración 8. Esquema fase a fase

7.4. Descripción del Trabajo de Aplicación

7.4.1. Descripción de la Empresa.

Publicar S.A. es una empresa enfocada en apoyar el desarrollo de la Pyme aumentando su visibilidad a través de publicidad multimedia, ofreciendo soluciones digitales y gráficas para la gestión de su empresa y entregando oportunidades y clientes reales para mover su negocio. Es por lo anterior que surgió la necesidad de realizar un análisis y ajuste del árbol de Latinoamérica del norte – LAN y la homologación del árbol de Latinoamérica del sur - LAS al de Latinoamérica del norte – LAN, con el fin de unificar los árboles y poder contar con una estructura de la información adecuada para su recuperación y así lograr que los clientes tengan más oportunidad de ser encontrados por los usuarios.

7.4.2. Diagnóstico.

Una vez verificado el árbol de Latinoamérica del Norte – LAN y los árboles de Latinoamérica del Sur – LAS, se pudieron identificar los siguientes puntos:

1. El árbol LAN está conformado por 13.702 registros.
2. Los árboles se encuentran conformados por tres niveles, los cuales son: primer nivel (*macrosecciones o categorías*) que hacen referencia a sectores económicos, segundo nivel (*rubros o subsecciones*) que segmenta los sectores económicos delimitando la actividad económica y el tercer nivel (*secciones o keywords*) que especifica la actividad económica o nicho de mercado del cliente.
3. El árbol LAN cuenta con 81 macrosecciones o categorías, 812 rubros o subsecciones y 9.172 secciones o keywords, de las cuales se hace necesario mover, unificar, renombrar y crear algunas de ellas para agrupar mejor los términos para la recuperación de la información.
4. Los árboles LAS son de los países de Perú y Chile, los cuales se deben homologar al árbol LAN teniendo en cuenta su estructura después de analizada y ajustada.
5. Los árboles LAS cuentan con los siguiente número de registros: Perú 283.530 y Chile 78.763.
6. El árbol LAS – Perú cuenta con 46 macrosecciones o categorías, 1.822 rubros o subsecciones y 135.669 secciones o keywords.
7. El árbol LAS – Chile cuenta con 41 macrosecciones o categorías, 3.443 rubros o subsecciones y 54.116 secciones o keywords.
8. Se precisa definir una estructura para el árbol definitivo la cual permita a Publicar aumentar las posibilidades de que sus clientes sean encontrados mediante su publicación en los medios multimedia y pueda agrupar la información de Latinoamérica del Norte con la de Latinoamérica del Sur.
9. Se identificaron casos como la duplicidad de secciones con diferentes macrosecciones y rubros, dando paso a la confusión del motor de búsqueda al momento de realizar la consulta de la información.
10. Se encontraron ID de secciones iguales con diferentes nombres, lo cual es un problema para poder encontrar lo que se necesita.
11. Se pueden encontrar nombres que no son los apropiados para que los usuarios al realizar la búsqueda puedan encontrar lo que necesitan.

12. Algunas macrosecciones, rubros y secciones deben ser unificadas porque se encuentran con nombres similares y esto desvía la búsqueda dejando algunos clientes por fuera de ella y sin posibilidad de tener visibilidad en el mercado multimedia.
13. Se debe realizar la identificación de aquellas secciones que se encuentran patrocinadas y cuantos clientes tienen estas ya sean pagos o gratuitos, esto para que el área encargada pueda definir el tiempo y los recursos que se deben emplear para el ajuste de los arboles LAN – LAS.
14. Al realizar la homologación se debe tener en cuenta que pueden surgir más ajustes dando que pueden aparecer nuevas secciones, rubros o macrosecciones, dependiendo de la concordancia de términos que se identifique en esta, dando paso así a renombrar algunas secciones para poder agrupar los términos y no generar duplicidad por tener varias palabras relacionadas.

7.4.3. Desarrollo de las Actividades.

Para el desarrollo del trabajo de aplicación se realizaron actividades tendientes a lograr los objetivos propuestos, estas actividades se mencionan a continuación en el orden en el cual se realizaron:

1. Definición de la estructura para el árbol final, la cual tiene las siguientes casillas:

ID	Macrosección	Descripción	ID	Rubros	Descripción	ID	Secciones
----	--------------	-------------	----	--------	-------------	----	-----------

2. Análisis del árbol LAN en sus tres niveles Macrosecciones, Rubros y Secciones, obteniendo como resultado la disminución en cada uno de ellos, optimizando la búsqueda en el portal web, tal como se puede evidenciar en la siguiente tabla:

Tabla 4. Datos finales del análisis del árbol LAN

Nivel	LAN		
	Iniciales	Finales	Diferencia
Macrosecciones	81	51	30
Rubros	812	536	276
Secciones	9.172	8.814	358

3. Ajuste del árbol LAN en sus tres niveles para obtener un árbol depurado y con la estructura adecuada para la correcta recuperación de la información.

Tabla 5. Ejemplo Base Inicial LAN

ID	Macroseccion	ID	Rubro	ID	Sección
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	2602	Fresas
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	14765	Fruta Deshidratada
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	2604	Frutas Congeladas
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	2605	Frutas Cristalizadas
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	2606	Frutas Exóticas
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	14973	Frutas y Verduras
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	2607	Fruterías
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	211299	Hortalizas, frutas u otros frutos conservados en vinagre
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	211354	Manzanas,peras,membrillos
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	5785	Piñas
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	4222	Plátano
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	7273	PULPA DE MANGO
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	7273	Pulpas de Fruta.
18017	ALIMENTOS Y BEBIDAS	20825	Frutas	211356	Uvas frescas o secas, incluidas las pasas
18017	ALIMENTOS Y BEBIDAS	20607	Harina	33206	Harina de Arroz
18017	ALIMENTOS Y BEBIDAS	20607	Harina	2769	Harina de Carne
18017	ALIMENTOS Y BEBIDAS	20607	Harina	2770	Harina de Pescado
18017	ALIMENTOS Y BEBIDAS	20607	Harina	6970	Harina de trigo y de morcajo

Tabla 6. Ejemplo de la base ajustada LAN

ID Macro	Macrosección	ID Rubro	Rubro	ID Sección	Sección
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	440	Arepas
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	4896	Pizza
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	4999	Sandwiches
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	5263	Tamales
14751	ALIMENTOS Y BEBIDAS	20731	Alimentos Dieteticos y Naturales	4346	Productos Naturales
14751	ALIMENTOS Y BEBIDAS	18303	Bebidas y Derivados Lacteos	2775	Helados
14751	ALIMENTOS Y BEBIDAS	18303	Bebidas y Derivados Lacteos	4344	Productos Lacteos
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	6272	Asaderos
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	274968	Carnes en General
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	1099	Carnes Frias
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	1098	Distribuidores Carnes
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	3246	Lechonas
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	274769	Lechonerias
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	4894	Mariscos y Pescados
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	3562	Mataderos
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	12559	Pescaderias
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	4897	Pollos Asados
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	4256	Pollos Crudos
14751	ALIMENTOS Y BEBIDAS	18304	Carnes, Pollos, Pescados y Mariscos	211517	Sushi

4. Generación del reporte de movimientos de rubros entre macrosecciones y de secciones entre rubros; unificación de macrosecciones, rubros y secciones; el de renombrar macrosecciones, rubros y secciones y la creación de macrosecciones, rubros y

secciones, para realizarlos en el portal web y así poder dejar listo el árbol LAN para la homologación de los árboles LAS.

Tabla 7. Ejemplo del reporte de ajustes

No.	Macrosección Inicial	ID	Acción	ID	Nombre Definitivo
1	TEXTOS Y LIBROS ESCOLARES	10083	Renombrar	10083	LIBROS Y LIBRERIAS
2	EQUIPOS PARA MEDICINA GENERAL	10135	Renombrar	10135	EQUIPOS Y SUMINISTROS PARA MEDICINA
3	EMPAQUES - VIDRIOS Y PLÁSTICOS	18128	Renombrar	18128	EMPAQUES, PAPEL, VIDRIOS Y PLASTICOS
4	TEXTILES Y PRENDAS DE VESTIR	18208	Renombrar	18208	CONFECCIONES, TEXTILES Y PRENDAS DE VESTIR
5	SALUD	18241	Renombrar	18241	SALUD Y BELLEZA
6	RECREACIÓN Y DIVERSIÓN	18247	Renombrar	18247	RECREACION, DIVERSION Y ENTRETENIMIENTO
7	EQUIPOS Y SUMINISTROS	18662	Renombrar	18662	EQUIPOS DE TECNOLOGIA Y SUMINISTROS INFORMATICOS
8	SERVICIOS	19629	Renombrar	19629	SERVICIOS DE ASESORIA, CONSULTORIA Y OUTSOURING
9	INDUSTRIA QUÍMICA	19650	Renombrar	19650	INDUSTRIA QUIMICA Y MINERALES BÁSICOS
10	OTRAS INDUSTRIAS	19654	Renombrar	19654	INDUSTRIAS MANUFACTURERAS
11	MAQUINARIA Y EQUIPO	19658	Renombrar	19658	MAQUINARIA Y EQUIPO (METALICOS Y EN GENERAL)
12	EQUIPOS Y SUMINISTROS PARA ESPECIALIDADES MÉDICAS	10132	Unificar	10135	EQUIPOS Y SUMINISTROS PARA MEDICINA
13	ALIMENTOS Y BEBIDAS	18017	Unificar	14751	ALIMENTOS Y BEBIDAS
14	ALIMENTOS Y BEBIDAS	19626	Unificar	14751	ALIMENTOS Y BEBIDAS
15	CONSTRUCCIÓN	31817	Unificar	18075	CONSTRUCCION (EDIFICACIONES, VIVIENDA Y OTROS)
16	EDUCACIÓN Y FORMACIÓN	46280	Unificar	18122	EDUCACION Y FORMACION
17	EMPAQUES Y PAPEL	31819	Unificar	18128	EMPAQUES, PAPEL, VIDRIOS Y PLASTICOS
18	EMPAQUES, ENVASES Y COMPONENTES	32476	Unificar	18128	EMPAQUES, PAPEL, VIDRIOS Y PLASTICOS

- La homologación de los árboles LAS se realizó en el siguiente orden: Perú y Chile, para este proceso se tuvo en cuenta el segundo nivel de LAS (Rubros) para ser homologado

por el tercer nivel de LAN (Secciones), dado que estos niveles son los comerciales de acuerdo a cada zona de Latinoamérica.

Tabla 8. Ejemplo de la homologación LAS a LAN

ID	RUBRO	Cuenta	ID Nuevo	Sección Definitiva LAN	Estado	Nombre Definitivo
940259	ASESORIAS TRIBUTARIAS Y CONTABLES	148	629	Asesorias Contables	Renombrar	Asesorias Tributarias y Contables
941356	CABOS Y JARCAS	3	936	Cabos	Renombrar	Cabos y Jarcas
943028	CENTROS DE DEPILACION	42	1710	Depilacion	Renombrar	Centros de Depilacion
943455	COLECTORES DE POLVO Y GASES	2	1394	Colectores de Polvo	Renombrar	Colectores de Polvo y Gases
941143	COLGADORES Y GANCHOS PARA ROPA	27	33444	Exhibidores y Colgadores para Ropa	Renombrar	Exhibidores, Colgadores y Ganchos para Ropa

- Identificar las secciones patrocinadas y la cantidad de clientes pagos y gratuitos que tienen.

Tabla 9. Ejemplo de secciones patrocinadas y cantidad de clientes

ID Macro	Macrosección	ID Rubro	Rubro	ID Sección	Sección	Patrocinadas	Pagos	Gratuitos
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	440	Arepas	440	44	451
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	4896	Pizza	4896	126	1971
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	4999	Sandwiches	4999	42	133
14751	ALIMENTOS Y BEBIDAS	18295	Alimentos Congelados, Refrigerados y Precocidos	5263	Tamales	5263	33	270

- Ajustes del árbol LAN – LAS definitivo para ser utilizado por Publicar en sus medios multimedia.

8. Recursos

8.1. Humanos

Para el desarrollo del trabajo de aplicación se contó con el apoyo de una Analista de Contenido Sintáctico quien verificaba el análisis realizado de los árboles para tener certeza que la base estaba quedando como se necesitaba y la Jefe de Producto Digital de Publicar quien aportó la información actualizada de los árboles y apoyo la solución de algunas inquietudes frente a la unificación de algunas macrosecciones, rubros y secciones, permitiendo así avanzar rápidamente con el análisis y ajuste del árbol LAN, así como la homologación de los árboles LAS al LAN.

8.2. Técnicos

En el desarrollo del trabajo de aplicación se empleó un equipo de cómputo portátil, con el cual se realizó el análisis, ajuste y homologación de los árboles de Latinoamérica del Norte y Sur LAN – LAS.

9. Resultados

Del desarrollo de este trabajo de aplicación se pudieron obtener los siguientes resultados:

- Una base de datos normalizada acorde a los parámetros dados por las normas y la empresa, logrando contar con esta para mejorar la visibilidad de los clientes en el portal web de Publicar.
- Homologación de los árboles LAS (Chile y Perú) a LAN de forma que se puedan mejorar las posibilidades de que las Pyme puedan aparecer en las búsquedas realizadas y explotar la base con un buen posicionamiento en Google.
- Poder contar con una base unificada para Latinoamérica y poder lograr su visión de “Ser líderes y aliados #1 en Publicidad Multimedia de la Pyme en América Latina”.
- Contar con una estructura adecuada para la futura creación de secciones, cumpliendo así con todos los requisitos de normalización ya identificados y planteados, después del análisis.
- Obtener palabras claves que permitan mejorar la identificación de las secciones por el motor de búsqueda al realizar la consulta.

10. Conclusiones

De la realización de este trabajo de aplicación se llegó a las siguientes conclusiones:

- La normalización de términos en una base tan grande como la del portal web de Publicar, beneficia totalmente a la empresa dado que podrá obtener mejores resultados en las búsquedas realizadas por los usuarios de este.
- Al realizar la homologación de términos se logra obtener una base unificada y con mayor posibilidad de brindar visibilidad a sus clientes durante las consultas en los motores de búsqueda de la web.
- Obtener una estructura para la información permite mejorar considerablemente la organización de la información y a su vez la posibilidad de identificar aquellas palabras o términos que no cumplen con la normalización realizada.
- Las palabras claves juegan un papel importante en una base como estas, ya que es utilizada en un portal web y mejoran la oportunidad de encontrar rápidamente lo que se necesita, sin perder tiempo revisando varias opciones.

11. Recomendaciones

Después del desarrollo de este trabajo de aplicación se recomienda tener en cuenta los siguientes puntos:

- Es necesario mantener la estructura definida para la creación de nuevas secciones lo cual permitirá tener un control más adecuado de la información recibida y retenida en el portal web de Publicar.
- Realizar la revisión periódica de la base de datos para poder identificar la posible duplicidad de información.
- Mejorar en algunos aspectos como la definición de los nombres para poder dar visibilidad a los clientes del portal, ya que esto permitirá atraer nuevos clientes y mejorar el servicio que se les brinda mediante el mercado multimedia.
- Conservar los niveles de información jerárquica para así lograr tener una cadena correcta de términos relacionados, brindando la posibilidad de agrupar adecuadamente la información que se tiene y la que puede ingresarse a futuro.
- Lograr tener la administración de la base controlada e involucrar pocos usuarios para no perder la estructura y prevenir la duplicidad de registros.

12. Referencias

Almazán, M. (2012). Herramientas para la Interoperabilidad y Normalización de datos en RI.
http://sedici.unlp.edu.ar/bitstream/handle/10915/25088/Documento_completo____.pdf?sequence=1

Born, M. (1999). La Terminología: Representación y Comunicación María Teresa Cabré
(Barcelona: Institut Universitari de Lingüística Aplicada.
http://onomazein.letras.uc.cl/Articulos/6/R1_Born.pdf

Castillo, J. (2010). Nueva propuesta evolutiva para el agrupamiento de documentos en
sistemas de información. https://ebuah.uah.es/dspace/bitstream/handle/10017/9551/Tesis-Jose_Luis_Castillo_Sequera.pdf?sequence=1&isAllowed=y

Castro, L. (2016). Análisis de la Información Análisis y Diseños de Sistemas de Información
II. Guía No.3. Análisis Estructurado. Universidad del Quindío.

Castro, L. (2016). Análisis de la Información Análisis y Diseños de Sistemas de Información
II. Guía No.2. Diseño de Salida de un Sistema. Universidad del Quindío.

Cruz, M (s.f.) Conceptos básicos de bases de datos.
<http://www.gridmorelos.uaem.mx/~mcruz/cursos/miic/bd1.pdf>

Gavilán, C. (2009). Lenguajes documentales Principales tipos de clasificación
Encabezamientos de materia, descriptores y tesauros.
<http://eprints.rclis.org/14817/1/lendoc.pdf>

Henao, M. (2014). Lingüística Documental. Guía No.3 La Estructura del Lenguaje.
Universidad del Quindío.

Henao, M. (2014). Semiótica. Guía No.1 La Ciencia Semiótica. Universidad del Quindío.

Laguens, J. (2006). Construcción de un tesoro especializado en hidrocarburos.

<http://revistas.um.es/analesdoc/article/viewFile/1391/1441>

Lancaster, F. (2002). El control del vocabulario en la recuperación de la información Capítulo 15: La influencia del vocabulario en el funcionamiento de un sistema de recuperación. 2da Edición. <https://bit.ly/2sht0ax>

Molina, M. (2011). Electronic Content Management Skills. <http://www.mariapinto.es/e-coms/busqueda-y-recuperacion-de-informacion/>

Palma, W. (2008). Adaptación de las directrices españolas y norteamericanas para la redacción y construcción de autoridades de materia en español para américa latina. <http://www.javeriana.edu.co/biblos/tesis/comunicacion/tesis63.pdf>

Parra, A. (2015). Lenguajes Documentales. Guía No.1 Lenguajes Documentales e Indización. Universidad del Quindío.

Parra, A. (2015). Lenguajes Documentales. Guía No.4 Tesoros. Universidad del Quindío.

Parra, A. (2015). Lenguajes Documentales. Guía No.2 Lenguaje de Análisis y Recuperación de la Información - LARI. Universidad del Quindío.

Pérez, Ch. (2014). Explotación de los corpórea textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento 3.2 Terminología descriptiva y prescriptiva (normalizadora). <http://elies.rediris.es/elies18/321.html>

Rubio, M. (s.f.). El Análisis Documental: Indización y Resumen en Bases de Datos Especializadas.

http://eprints.rclis.org/6015/1/An%C3%A1lisis_documental_indizaci%C3%B3n_y_resumen.pdf

Santamaría, I. La terminología: definición, funciones y aplicaciones.

https://rua.ua.es/dspace/bitstream/10045/12770/8/La_terminolog%C3%ADa.pdf

Torres, J. (2016). Capítulo 1: Conceptos Básicos de Bases de Datos.

<https://slidex.tips/download/capitulo-1-conceptos-basicos-de-bases-de-datos>

Triana, J. (2015). Construcción de un tesoro especializado en comunicación social para el contexto colombiano.

http://repository.lasalle.edu.co/bitstream/handle/10185/17393/33081208_2015.pdf?sequence=1

Universidad de la Punta. Escuela Pública Digital. (s.f.) Unidad 1. Sistemas Gestores de Bases de Datos.

http://contenidosdigitales.ulp.edu.ar/exe/computacion2/sistemas_gestores_de_bases_de_datos.html

Vargas, N., Paredes, L. (2010). Normalización del lenguaje documental en antropología andina boliviana, en base a un prospecto de microtesoro especializado.

<http://repositorio.umsa.bo/bitstream/handle/123456789/6624/T-%202371.pdf?sequence=1&isAllowed=y>

Vilares, J. Barcala, M. Alonso, M. (s.f.). Normalización de términos multpalabra mediante pares de dependencia sintáctica.

https://rua.ua.es/dspace/bitstream/10045/1794/1/PLN_27_14.pdf

Wüster, E. (2010). Introducción a la teoría general de la terminología y a la lexicografía terminológica. 3.2. Relaciones entre conceptos (sistemas de conceptos). <https://bit.ly/2Im0Lla>

13. Anexo (Carta de Aceptación Trabajo de Aplicación)



Bogotá, 13 de Marzo de 2018

Señores
Programa CIDBA
Universidad del Quindío
Ciudad

Asunto: Autorización trabajo de aplicación.

De acuerdo con la propuesta de trabajo de aplicación presentada por la estudiante Lizeth Giovana Rodríguez Calderón identificada con cedula de ciudadanía 1.018.415.277 de Bogotá, que tiene como título *"Normalización y homologación de los árboles de información de Latinoamérica del Norte y el Sur de la empresa PUBLICAR S.A."*, la cual se llevara a cabo entre las fechas del 2 de marzo y el 27 de abril del presente año,

Teniendo en cuenta lo anterior se autoriza el desarrollo del trabajo de aplicación que tiene como objetivo *"Realizar la normalización y homologación de la información recibida o retenida por el portal web de publicar, estructurando jerárquica y semánticamente los términos de búsqueda o palabras clave, permitiendo una óptima recuperación de información comercial requerida por clientes, proveedores y usuarios en general de los países latinoamericanos donde opera PUBLICAR S.A."* y del cual se obtendrá unos productos finales que serán el árbol LAN analizado y normalizado, los árboles LAS homologados a la estructura del árbol LAN y los reportes de los ajustes que se deben realizar en el sistema de acuerdo con el análisis y la normalización de términos.

Atentamente,



NTT: 900.432.040 - 3 Resolución: 320001036093
Rango: 99 - 1000 Fecha: 2013-07-11

Juan Eduardo Gómez
Gerente General
MENTOR 360 S.A.S.
Celular: 300-5165444
Mail: juan.gomez@mentor360.net

14. Anexo (Consentimiento Informado)



UNIVERSIDAD DEL QUINDÍO
FACULTAD DE CIENCIAS HUMANAS Y BELLAS ARTES
CIENCIA DE LA INFORMACIÓN Y LA DOCUMENTACIÓN, BIBLIOTECOLOGÍA Y
ARCHIVÍSTICA.
COMITÉ TRABAJOS DE GRADO

CONSENTIMIENTO INFORMADO

Yo **JUAN EDUARDO GOMEZ FERNANDEZ** identificado con la C.C. 92'535.932 como Representante legal o Director de la Institución o entidad **MENTOR 360 S.A.S.** autorizo la realización del Trabajo de aplicación denominado **"Normalización y homologación de los árboles de información de Latinoamérica del Norte y el Sur de la empresa PUBLICAR S.A."** Así mismo, acepto verificar los resultados del trabajo ya que he sido informado (a) acerca de sus objetivos, cronograma y lo que se espera del mismo.

Entiendo que la estudiante de CIDBA encargada de la intervención, puede detener la aplicación o la participación de esta entidad en cualquier momento sin mi consentimiento. Así mismo, tengo derecho a retirarme en cualquier momento.

Por el presente autorizo a la estudiante de CIDBA a publicar la información obtenida como resultado de la participación en el trabajo de aplicación, en la sustentación u otros medios legales y de permitirles revisar el proyecto guardando la debida CONFIDENCIALIDAD de mí y de esta entidad.



Firma del Responsable:

Juan Gomez

Firma de estudiante (s) CIDBA:

Urieth Rodriguez

Fecha: 13 de Marzo de 2018

Resolución N° 008430 de 1993
Normas éticas para la investigación

15. Anexo (Carta Finalización Trabajo de Aplicación)



Bogotá, 03 de Mayo de 2018

Señores
Programa CIDBA
Universidad del Quindío
Ciudad

Asunto: Finalización trabajo de aplicación.

De acuerdo con el desarrollo del trabajo de aplicación de la estudiante Lizeth Giovana Rodríguez Calderón, identificada con cedula de ciudadanía 1.018.415.277 de Bogotá, que tiene como título *"Proyecto Llave en mano - Normalización y homologación de los árboles de información de Región Norte y Sur de la empresa PUBLICAR PUBLICIDAD Y MULTIMEDIA S.A.S."*, el cual se llevó a cabo entre las fechas del 2 de marzo al 27 de abril del presente año, con una dedicación de cuatro (4) horas diarias para un total de 136 horas.

Teniendo en cuenta lo anterior es importante mencionar que se cumplió con el objetivo general del trabajo el cual consistió en *"Realizar la normalización y homologación de la información recibida o retenida por el portal PaginasAmarillas.com, estructurando jerárquica y semánticamente los términos de búsqueda o palabras clave, permitiendo una óptima recuperación de información comercial requerida por clientes y usuarios en general de los países latinoamericanos donde opera PUBLICAR."*, obteniendo los resultados esperados, obteniendo la normalización y homologación de los árboles de categorías LAN – LAS.

Atentamente,



MENTOR 360 R T.I. SOLUCIONES
NIT: 3001422040 - Resolución: 320001036093
Rango: 99 - 1000 Fecha: 2013-07-11

Juan Eduardo Gómez
Gerente General
MENTOR 360 S.A.S.
Celular: 300-5165444
Mail: juan.gomez@mentor360.net

16. Anexo (Informe Resultados Trabajo de Aplicación)



UNIVERSIDAD DEL QUINDÍO



FACULTAD DE CIENCIAS HUMANAS Y BELLAS ARTES

CIENCIA DE LA INFORMACIÓN Y LA DOCUMENTACIÓN, BIBLIOTECOLOGÍA Y
ARCHIVÍSTICA.

COMITÉ TRABAJOS DE GRADO

INFORME DE RESULTADOS DEL TRABAJO DE APLICACIÓN.

ESTUDIANTE	Lizeth Giovana Rodríguez Calderón
INSTITUCIÓN O EMPRESA	Mentor 360
DIRECCIÓN	Carrera 46 No. 123-30
EVALUADOR DEL PROCESO	Sandy Salazar Leal
CARGO	Jefe de Producto Digital
TELÉFONO DE CONTACTO	315 259 4218
APLICACIÓN - FECHA	2 marzo al 27 abril 2018
ASESOR	Angélica María Ramírez Agudelo

1. Concepto de quien evalúa el proceso con respecto al desempeño del estudiante:

El desempeño de Lizeth en el proyecto fue excelente dado a su gran compromiso y dedicación en los resultados del análisis, en el cual demuestra su amplio conocimiento y criterio sobre el análisis requerido. El resultado fue el esperado.

2. ¿Qué objetivos se cumplieron en la aplicación del trabajo realizado?

El objetivo principal fue la Normalización y homologación del árbol de categorías de Pádel (páginas amarillas) .com para los países de Latinoamérica norte y sur, donde se estructuró y jerarquiza la información la cual permite una optimización y un correcto uso para el usuario, como para la tienda comercial.



3. Qué impacto tuvo el trabajo de aplicación en su Institución o Empresa:

Este proyecto le permite a publicar, por una parte, la unificación de los árboles de información para el portal, en presencia en 3 países, esta Ayuda a controlar las categorías (comercializadas publicidad), como la indexación a nivel de motores de búsqueda dado a que el modelo anterior no permitía exhibir la base de información con su posicionamiento en Google.

4. Observaciones del estudiante (Autoevaluación):

Se desarrolló el trabajo de aplicación para la normalización de la base de datos de información retenida por el portal web de Publicar S.A., para lo cual se tuvieron sesiones de seguimiento, identificando así los puntos a mejorar e interactuando con las personas encargadas de la supervisión. Por otra parte mediante la realización de este fortalecimiento concimientos acerca de la normalización del lenguaje, además de tener la oportunidad de mejorar y fortalecer habilidades como la comunicación, toma de decisiones y recopilación de información para entregar un producto adecuado y funcional para la empresa.


FIRMA DEL EVALUADOR


FIRMA DEL ESTUDIANTE

17. Anexo (Control Tiempo Trabajo de Aplicación)



FORMATO CONTROL TIEMPO DE TRABAJO DE APLICACIÓN

No.	Nombre	Fecha	Hora Inicio/ Hora Fin	Total Horas	Supervisor y Cargo	Firma
1	Lizeth Giovana Rodríguez Calderón	1/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
2	Lizeth Giovana Rodríguez Calderón	2/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
3	Lizeth Giovana Rodríguez Calderón	5/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
4	Lizeth Giovana Rodríguez Calderón	6/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
5	Lizeth Giovana Rodríguez Calderón	7/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
6	Lizeth Giovana Rodríguez Calderón	8/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
7	Lizeth Giovana Rodríguez Calderón	9/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
8	Lizeth Giovana Rodríguez Calderón	12/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
9	Lizeth Giovana Rodríguez Calderón	13/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
10	Lizeth Giovana Rodríguez Calderón	14/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
11	Lizeth Giovana Rodríguez Calderón	15/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
12	Lizeth Giovana Rodríguez Calderón	16/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
13	Lizeth Giovana Rodríguez Calderón	20/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
14	Lizeth Giovana Rodríguez Calderón	21/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
15	Lizeth Giovana Rodríguez Calderón	22/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar

www.mentor360.net • Carrera 45 No 123 - 30 Of 204 • Teléfono: 5121962 • email: contacto@mentor360.net




No.	Nombre	Fecha	Hora Inicio/ Hora Fin	Total Horas	Supervisor y Cargo	Firma
16	Lizeth Giovana Rodríguez Calderón	23/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
17	Lizeth Giovana Rodríguez Calderón	26/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
18	Lizeth Giovana Rodríguez Calderón	27/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
19	Lizeth Giovana Rodríguez Calderón	28/03/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
20	Lizeth Giovana Rodríguez Calderón	3/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
21	Lizeth Giovana Rodríguez Calderón	10/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
22	Lizeth Giovana Rodríguez Calderón	11/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
23	Lizeth Giovana Rodríguez Calderón	12/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
24	Lizeth Giovana Rodríguez Calderón	13/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
25	Lizeth Giovana Rodríguez Calderón	16/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
26	Lizeth Giovana Rodríguez Calderón	17/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
27	Lizeth Giovana Rodríguez Calderón	18/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
28	Lizeth Giovana Rodríguez Calderón	19/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
29	Lizeth Giovana Rodríguez Calderón	20/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
30	Lizeth Giovana Rodríguez Calderón	23/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
31	Lizeth Giovana Rodríguez Calderón	24/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar
32	Lizeth Giovana Rodríguez Calderón	25/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	Sandy Salazar

www.mentor360.net • Carrera 45 No 123 - 30 Of 204 • Teléfono: 5121962 • email: contacto@mentor360.net



No.	Nombre	Fecha	Hora Inicio/ Hora Fin	Total Horas	Supervisor y Cargo	Firma
33	Lizeth Giovana Rodríguez Calderón	26/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	<i>Sandy Salazar</i>
34	Lizeth Giovana Rodríguez Calderón	27/04/2018	13:00 a 17:00	4 horas	Sandy Salazar Jefe de Producto Digital	<i>Sandy Salazar</i>
				136		

Juan Eduardo Gómez

IMPRESO POR: 300-5165444 - T. Resolución: 220257038093
 Fecha: 26/04/2018 11:11
Juan Eduardo Gómez
 Gerente General
 MENTOR 360 S.A.S.
 Celular: 300-5165444
 Mail: juan.gomez@mentor360.net

18. Anexo (Registros de Asistencia Reuniones)



FORMATO CONTROL DE ASISTENCIA

Evento: Reunión Seguimiento PACOM				Fecha: 26 marzo 2018
No.	Nombre	Entidad	Teléfono/Correo	Firma
1				
2	Irish Giovana Rodríguez	Mentor 360	irishgiovana@mentor360.net	Irish Rodríguez
3	Sandy Salazar	Publicar	Sandy.Salazar@Publicar	Sandy Salazar
4	Volando Serrano	Publicar	Volando.Serrano@Publicar	Volando Serrano
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

www.mentor360.net • Carrera 46 No 529 - 50 Of 204 • Teléfono: 6121982 • email: contacto@mentor360.net



FORMATO CONTROL DE ASISTENCIA

Evento: Reunión Seguimiento PACOM				Fecha: 5/marzo 2018
No.	Nombre	Entidad	Teléfono/Correo	Firma
1				
2	Irish Giovana Rodríguez	Mentor 360	irishgiovana@mentor360.net	Irish Rodríguez
3	Sandy Salazar	Publicar	Sandy.Salazar@Publicar	Sandy Salazar
4	Volando Serrano	Publicar	Volando.Serrano@Publicar	Volando Serrano
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

www.mentor360.net • Carrera 46 No 529 - 50 Of 204 • Teléfono: 6121982 • email: contacto@mentor360.net



FORMATO CONTROL DE ASISTENCIA

Evento: Reunión seguimiento Pacom			Fecha: 12 marzo 2018	
No.	Nombre	Entidad	Teléfono /Correo	Firma
1				
2	Irish Givans Rodríguez	Mentor 360	IrishGivansRodriguez@gmail.com	Irish Rodríguez
3	Sandy Salazar	Publicar	Sandy.Salazar@publicar	Sandy Salazar
4	Rolando Serrano	Publicar	Rolando.Serrano@publicar	Rolando Serrano
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

www.mentor360.net • Carrera 46 No. 523 - 30 Of 204 • Teléfono: 5121362 • email: contacto@mentor360.net



FORMATO CONTROL DE ASISTENCIA

Evento: Reunión Seguimiento Pacom			Fecha: 2 abril 2018	
No.	Nombre	Entidad	Teléfono /Correo	Firma
1				
2	Irish Givans Rodríguez	Mentor 360	IrishGivansRodriguez@gmail.com	Irish Rodríguez
3	Sandy Salazar	Publicar	Sandy.Salazar@publicar	Sandy Salazar
4	Rolando Serrano	Publicar	Rolando.Serrano@publicar	Rolando Serrano
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

www.mentor360.net • Carrera 46 No. 523 - 30 Of 204 • Teléfono: 5121362 • email: contacto@mentor360.net



FORMATO CONTROL DE ASISTENCIA

Evento: Reunión Seguimiento Pacem.			Fecha: 09 Abril 2018	
No.	Nombre	Entidad	Teléfono /Correo	Firma
1				
2	Lizeth Girona Rodríguez	Mentor 360	lizeth.girona@mentor360.net	Lizeth Rodríguez
3	Sandy Salazar	Publicar	Sandy.Salazar@Publicar	Sandy Salazar
4	Rolando Serrano	Publicar	rolando.serrano@Publicar	Rolando Serrano
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

www.mentor360.net • Carrera 42 No 123 - 30 Of 204 • Teléfono: 0121982 • email: contacto@mentor360.net



FORMATO CONTROL DE ASISTENCIA

Evento: Reunión Seguimiento Pacem.			Fecha: 16 abril 2018	
No.	Nombre	Entidad	Teléfono /Correo	Firma
1				
2	Lizeth Girona Rodríguez	Mentor 360	lizeth.girona@mentor360.net	Lizeth Rodríguez
3	Sandy Salazar	Publicar	Sandy.Salazar@Publicar	Sandy Salazar
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				

www.mentor360.net • Carrera 42 No 123 - 30 Of 204 • Teléfono: 0121982 • email: contacto@mentor360.net



FORMATO CONTROL DE ASISTENCIA

Evento: Reunión Seguimiento Pacan		Fecha: 23 abril 2018		
No.	Nombre	Entidad	Teléfono /Correo	Firma
1				
2	Urieth Gausso Rodríguez	Mentor 360	investigacion@mentor360.net	Urieth Rodríguez
3	Sandy Salazar	Publicar	Sandy Salazar Rodríguez	Sandy Salazar
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				



UNIVERSIDAD
DEL QUINDÍO

Por una Universidad
PERTINENTE CREATIVA INTEGRADORA

