CONTROL DE AUTORIDADES Y ESTANDARIZACIÓN DE METADATOS EN LAS MATERIAS Y EN LOS AUTORES QUE SE HAYAN IDENTIFICADO DEL REPOSITORIO INSTITUCIONAL DE LA UNIVERSIDAD DE ANTIOQUIA

YANNY MARCELA VACA PALACIOS



UNIVERSIDAD DE ANTIOQUIA
ESCUELA INTERAMERICANA DE BIBLIOTECOLOGÍA
MEDELLÍN
2018

CONTROL DE AUTORIDADES Y ESTANDARIZACIÓN DE METADATOS EN LAS MATERIAS Y EN LOS AUTORES QUE SE HAYAN IDENTIFICADO DEL REPOSITORIO INSTITUCIONAL DE LA UNIVERSIDAD DE ANTIOQUIA

YANNY MARCELA VACA PALACIOS

ASESOR INTERNO
MAURICIO ZAPATA OSORIO

UNIVERSIDAD DE ANTIOQUIA
1803
Escuela Interamericana de Bibliotecología

ASESOR EXTERNO
LINA MARÍA ZAPATA CASAS

INFORME DE PRÁCTICA PARA OPTAR EL TÍTULO DE BIBLIOTECÓLOGA

UNIVERSIDAD DE ANTIOQUIA
ESCUELA INTERAMERICANA DE BIBLIOTECOLOGÍA
MEDELLÍN
07/06/2018

Agradecimientos

Primero quiero darle las gracias a Dios porque gracias a Él pude estar en esta ciudad cumpliendo mis sueños.

A mis padres Amanda Palacios y Luis Eduardo Vacca por hacer posible todo lo que he querido y confiar en mí.

A mi asesor de práctica Mauricio Zapata por tantas enseñanzas brindadas para que yo pudiera lograr este trabajo.

Gracias.

CONTENIDO

Presentación 7
1. Objetivo general 8
1.1. Objetivos específicos 8
2. Marco teórico y conceptual 9
3. Marco contextual 13
4. Metodología 15
5. Desarrollo de actividades 16
6. Resultados 21
7. Conclusiones y recomendaciones 27
8. Referencias 28

ÍNDICE DE TABLAS

Tabla 1. Códigos handle del repositorio 18
Tabla 2: Resultados control de autoridades 22
Tabla 3. Registros validados 24
Tabla 4. Fuentes de validación 25

ÍNDICE DE GRÁFICOS

Gráfico 1. Representación de materias 22
Gráfico 2. Estandarización lenguaje controlado 23
Gráfico 3. Representación de autores 24
Gráfico 4. Porcentaje de fuentes de validación 26

ÍNDICE DE ANEXOS

Anexo 1. Validación de materias 29
Anexo 2. Validación de autores 30
Anexo 3. Listado de materias 31
Anexo 4. Tesauro de materias construido 31
Anexo 5. Nombres que se logran organizar 32
Anexo 6. Otros nombres 32
Anexo 7. Nombres organizados 33
Anexo 8. Términos relacionados por materias unificados 33
Anexo 9. Otros términos relacionados por materias 34
Anexo 10. Autores 34
Anexo 11. Identificación de los autores 34
Anexo 12. Convenciones descritas para autores 35
Anexo 13. Autores descritos con el rol en los registros 35
Anexo 14. Autor organizado en metadatos de notas 35
Anexo 15. Otros autores descritos con los roles 36
Anexo 16. Autores organizados en metadatos de notas 36
Anexo 17. Otros autores con roles descritos 36
Anexo 18. Varios nombres de autores 37
Anexo 19. Nombres de autores unificados 37

Resumen

El presente informe corresponde a la Práctica académica de bibliotecología realizada en el Repositorio Institucional de la Universidad de Antioquia, cuyo objetivo principal, permitió contribuir al mejoramiento de registros de información y del mismo repositorio. Las actividades se concentraron en la contextualización de los repositorios y en la estandarización y normalización de ***Autores*** y ***Materias***, mostrando mejores conjuntos de ítems asociados a términos específicos, los cuáles, mejoran a su vez, índices de recuperación y visibilidad de contenidos. Todo se hizo por medio de un proceso general de control de calidad.

Palabras Claves: Repositorio institucional, Control de autoridades, Estandarización de metadatos, Lenguaje controlado.

Presentación

Para el Repositorio Institucional del Sistema de Bibliotecas de la Universidad de Antioquia se constituye el control de autoridades como un proceso de suma importancia para la utilización de formas normalizadas de las materias y los autores en los puntos de acceso de los registros, esto hace que se tenga una mejor organización de los documentos, logren relacionarse entre sí con la finalidad de que sean identificados y recuperados y, a su vez, contribuya a que el usuario que hace la búsqueda pueda adquirir un documento mucho más organizado y encuentre de alguna manera siempre la misma forma de encontrar registros que además son validados por herramientas de gestión bibliotecarias para la catalogación y la normalización de la información.

Es importante también, el desarrollo de un tesauro gráfico en el Repositorio Institucional, ya que se convierte en una herramienta útil y de fácil acceso para la información registrada, esto con el fin de que se tenga un control de términos que pueden ayudar también a la gestión y la transferencia de la información.

Es un proyecto que se realizó para contribuir al mejoramiento del repositorio institucional de la Universidad de Antioquia, en el que se establece un proceso de control de calidad, teniendo en cuenta una etapa inicial de control de autoridades de materias en el que se logra la normalización, estandarización de metadatos, indización de los documentos que corresponde a la producción y un lenguaje controlado para facilitar el acceso y la recuperación de la información.

Y en la segunda etapa se hizo un control de autoridades de autores en el que se identificaron los nombres de los autores y se la validación de los mismos, en el que se desarrolló el mismo proceso aplicado a las materias.

1. Objetivo general

Establecer un control de autoridades y estandarización de metadatos en las materias y en los autores que se hayan identificado del Repositorio Institucional de la Universidad de Antioquia.

1.1. Objetivos específicos

- Revisar el manual de las políticas para el análisis de la información en el que ayuda a establecer de qué manera se va a tener el contenido de los documentos del repositorio.
- Avanzar en la metodología para la construcción de atributos en los lenguajes documentales que corresponden a las materias según tendencias que faciliten el acceso y recuperación de la información contenida en bases de datos y repositorios institucionales.
- Unificar descriptores desde el lenguaje controlado
- Validar el lenguaje controlado desde hermanitas con LEMB Digital, Tesauros, entre otras fuentes usadas.
- Realizar un control de autoridades con los autores que permita la estandarización para la búsqueda, la identificación y la recuperación de la producción del mismo, usando herramientas como VIAF, Cvlac Colciencias, catálogo OLIB, entre otras.

2. Marco teórico y conceptual

Repositorio institucional: Para Crow, (2002) lo define de forma muy escueta como una colección digital que agrupa y preserva la producción intelectual de una o varias universidades. Establece un objeto de aplicación muy amplio, el conjunto de la producción intelectual de la institución, en el que cabría más allá de los resultados de investigación y objetos de aprendizaje cualquier otro producto de carácter cultural producido por la institución.

Para Lynch, (2003) Define un repositorio institucional como un conjunto de servicios que una universidad ofrece a los miembros de su comunidad para la gestión y diseminación de los materiales digitales creados por la institución y sus miembros. Destacable la consideración de servicio que se da al repositorio, como algo dinámico al servicio de la comunidad, más allá de un depósito pasivo de documentos digitales.

Para McDowell, (2007) considera que un repositorio debe cumplir al menos las siguientes características:

- Es un servicio institucional abierto a toda la comunidad universitaria y a todo tipo de temáticas.
- Su objetivo debe ser reunir, preservar y dar acceso a, entre otras cosas, la producción de los investigadores y docentes en múltiple formatos. Se excluiría cualquier repositorio que ponga límites a esta producción, por ejemplo, sólo objetos de aprendizaje o sólo de tesis, etc.
- Debe recibir contenidos de forma activa bien a través de un formulario web o simplemente a través de correo electrónico.

Lenguaje de Indización: La operación de indizar consiste en retener una o más nociones que representan el contenido de un documento o los conceptos de una búsqueda bibliográfica. Se entiende por lenguaje de indización o lenguaje documental al conjunto controlado de términos o su representación simbólica, necesaria para poder comunicarnos con la información contenida. En un lenguaje controlado de indización los términos pertenecen a dos categorías (Estrada Villacís, 2015)

- Términos preferentes o descriptores o términos controlados.- Son términos que se utilizan de forma conveniente para representar conceptos precisos y unívocos en el proceso de indización y búsqueda de documentos.
- Términos no preferentes o no descriptores.- Se asignan a los documentos cuando indizamos, pero pueden proporcionar puntos de acceso en los índices o lenguajes impresos y tesauros y dirigir al usuario a la selección de descriptores a través de los términos sinónimos.

Así, la indización puede realizarse sobre la base de:

El texto del resumen

El título del documento

Términos controlados

Términos no controlados (términos libres que pertenecen al lenguaje natural y que se encuentran al arbitrio del profesional)

Tesauros: Una base de datos puede tener simplemente un vocabulario controlado o lenguaje de indización, o bien un tesauro. Éste supone que se señalan las relaciones conceptuales entre los descriptores, y la existencia de un vocabulario de entrada formado por los términos sinónimos o casi sinónimos de los descriptores (Estrada Villacís, 2015)

El tesauro reduce las ambigüedades entre términos, indica cómo puede haberse indizado un tema, alivia el problema de identificar todos los posibles términos que se utilizarán en las búsquedas, y proporciona la base para efectuar búsquedas al nivel de generalidad deseado.

Lista de encabezamientos de materia: Es un lenguaje documental controlado que recoge ordenadas alfabéticamente todas las materias o temas susceptibles de representar el contenido de un fondo documental. Es una lista, un léxico de palabras y frases uniformes que se usan para indizar la materia del documento (Juncá, 2010).

Control de autoridades: Es el proceso de unificar, mediante la utilización de una forma normalizada, los puntos de acceso de los catálogos automatizados y mostrar además las relaciones entre los distintos puntos de acceso. Es decir, supone la normalización de los nombres de personas, entidades, títulos uniformes o materias, que pueden constituir el punto de acceso principal o los secundarios de un catálogo automatizado. Su finalidad es facilitar la identificación y la recuperación de los documentos almacenados, evitando las confusiones a que se pueden prestar los homónimos, sinónimos o la variedad de nombres con los que puede ser denominado una persona, entidad, obra, tema o concepto (Nieto, 2013)

OPAC: (Online Public Access Catalogue) es el catálogo en línea en el que puedes consultar la existencia y ubicación de los libros, revistas, videos y demás material bibliográfico físico que se tiene en el Sistema de Bibliotecas y los Centros de Documentación de la Universidad. (Sistema de Bibliotecas, 2017).

VIAF: Es un proyecto internacional, liderado por la Library of Congress, la OCLC, y las bibliotecas nacionales de Francia y Alemania que combina registros de autoridad de varias bibliotecas nacionales. Identifica los registros de autoridad que corresponden al mismo autor, y los presenta en un único registro (cluster), que suma los registros individuales de todas las bibliotecas. En este momento, los países participantes son: EEUU, Francia, Alemania, Suecia, Egipto (que aporta registros en alfabeto árabe), Vaticano, Australia, República Checa, Israel (que aporta registros en alfabetos latino, eslavo y hebreo), Italia, Portugal, Suiza, República Checa, Polonia, Hungría, Japón, el Instituto Getty y la Biblioteca Nacional

de España. En esta primera fase, se han integrado los registros de autoridad de persona. En VIAF únicamente se accede al contenido de las autoridades, no se puede acceder a registros bibliográfico. (Biblioteca Nacional de España, 2010).

Scopus: Es la mayor base de datos de citas y resúmenes de literatura revisada por pares: revistas científicas, libros y actas de congresos. Scopus cuenta con herramientas inteligentes para rastrear, analizar y visualizar la investigación, ofreciendo una visión general de la producción mundial de investigación en los campos de ciencia, tecnología, medicina, ciencias sociales y artes y humanidades (Elsevier, 2018).

A medida que los trabajos de investigación se vuelven más globales, interdisciplinarios y colaborativos, es necesario asegurarse de no pasar por alto investigaciones cruciales realizadas alrededor del mundo.

WorldCat: Es la base de datos de información sobre colecciones bibliotecarias más completa del mundo**.** OCLC entrega calidad, capacidad de identificación y valor. Las bibliotecas miembro proporcionan la base. (OCLC, 2018).

Es una demostración de la creatividad e innovación del personal de OCLC y de miles de bibliotecarios. Única por su tamaño y sin igual por la calidad de sus datos, WorldCat facilita la identificación y el acceso a las colecciones bibliotecarias desde cualquier parte del mundo.

CvLAC y GrupLAC: Son herramientas informáticas que actualizan el sistema de bodega de datos (data warehouse) de Colciencias. Son servicios permanentes de hoja de vida en el cual los investigadores, innovadores, académicos, expertos y grupos que así lo deseen pueden actualizar su información para los usos del Sistema Nacional de Ciencia y Tecnología (SNCTeI) (Colciencias, s.f.)

Es el instrumento o formato electrónico de hoja de vida básico utilizado por Colciencias para:

• El registro permanente de hojas de vida de investigadores, innovadores, académicos y expertos del SNCTeI. El registro de www.colciencias.gov.co 3 hojas de vida de candidatos a ingresar al servicio de información de pares evaluadores reconocidos del SNCTeI. Las convocatorias para reconocimiento y medición de Grupos Colombianos de Investigación Científica o Tecnológica.

• Convocatorias de becas para formación de recurso humano, apoyo a programas doctorales a nivel nacional y jóvenes investigadores.

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

ResearchGate: Es una red social científica para científicos e investigadores cuya misión es conectar a investigadores y facilitarles compartir y acceder a sus publicaciones, conocimientos y experiencia (Moya Orozco, 2016).
Permite:
Compartir tus publicaciones y ganar visibilidad.
Conectar y colaborar con otros investigadores.
Obtener estadísticas y métricas sobre tu perfil y tus publicaciones.
Resolver cuestiones de forma colaborativa.

Google Académico o Google Scholar: En su versión en inglés es el buscador de Google especializado en literatura científica o académica, y te permite localizar de forma sencilla un montón de información que te resultará de gran ayuda a la hora de hacer tus trabajos para el colegio, el instituto o la universidad (Muñoz de Frutos, 2017).

3. Marco contextual

En la página de la Universidad de Antioquia en el Sistema de Bibliotecas, el repositorio institucional se concibe como una iniciativa que surge desde el mismo Sistema con el fin de que sea una herramienta de acceso libre, que se visibilice, fomento de la información, uso de las nuevas tecnologías y contribuya a la preservación de los materiales bibliográficos que corresponden a la institución, y más aún, las producciones que surjan de la misma.

El repositorio Institucional de la Universidad de Antioquia, fue creado en el 2002, el cual tiene diecisiete (17) años de existencia en coordinación con el Sistema de Bibliotecas del Alma Mater, con el fin de albergar toda la producción científica, académica, institucional y cultural. Se cuenta con siete (7) mil documentos, los cuales están distribuidos en nueve (9) comunidades (colecciones), divididos entre artículos, libros, trabajos de grado, materiales patrimoniales.

Desde el 2017, se concibe como una fuente institucional de información, y debido a su antigüedad no se ha hecho oficialmente un control de calidad en los procesos de flujos de trabajo, en los que se cuenta con la etapa de control de autoridad y de estandarización de descriptores, por ello se identifica la necesidad de examinar el estado actual de desarrollo de control de autoridades y estandarización de metadatos, esto con el fin, de que se contribuya en parte a lo que ya se tiene dentro del repositorio que se ha venido trabajando y darle continuidad, teniendo en cuenta que la primera etapa es de control de autoridades de materias con el uso de los tesauros de la Unesco, LEMB, de ciencias interdisciplinares que son los que ayudan a la validación de términos y a incrementar algunos que no se tengan y para la segunda etapa con los autores desarrollarlo de la misma manera y tratando de unificar de alguna manera todos los registros que constituyan al autor pero está escrito de una manera diferente.

El repositorio hace parte de la red de instituciones indizadas en SNACC- Sistema Nacional de Acceso Abierto en Colombia, que es una iniciativa de país impulsada desde el sector público y privado con el propósito de promover el acceso al conocimiento científico producido por distintas instituciones colombianas, y de fortalecer su visibilidad tanto a en el escenario nacional como en el internacional. Desde aquí, se alcanza un mayor nivel de divulgación de la producción académica, científica, cultural y patrimonial que hace parte del repositorio institucional.

Debido a que el repositorio institucional tiene como objetivo principal visibilizar la información, también se requiere que esta información que se muestra esté organizada de manera adecuada. El Repositorio Institucional de la Universidad de Antioquia cuenta con un comité editorial responsable de recomendar el tipo de trabajos que deben incluirse en el repositorio institucional, de acuerdo a las políticas y criterios establecidos previamente.

De esta manera, se evidencia un proceso de crecimiento a nivel tecnológico que es el que se adopta con el uso de las bibliotecas digitales, en el que se hace necesario que se cumplan con diversos esquemas de organización y control en el repositorio institucional; este se constituye en la carta de entrada para el conocimiento de la institución en la que se desee conocer y de igual manera, no solo en un servicio, sino en un servicio que es de y para la comunidad.

Dentro del desarrollo de este proyecto se establece que para el Sistema de Bibliotecas es muy importante, ya que se busca un mejor acceso a la información, en la que se denote que hay una organización en la información registrada que corresponden al repositorio institucional y que ayude a describir y representar su contenido. Además, porque no es solo necesario para el usuario sino para el analista, en el que se avanza en un proceso de continuidad y se desarrolla de una manera más cómoda, eficaz y efectiva en la recuperación de la información. Es un proyecto que se estableció con el fin de que se tenga un mejoramiento en el proceso de indización y de alguna manera contribuir al análisis, organización, disposición y estructuración de la información en el que se permite al usuario una mayor visibilidad y facilidad en el acceso.

4. Metodología

Para lograr el objetivo de este proyecto, se trata de establecer parámetros incurran en el mejoramiento del repositorio institucional.

En primera medida y para la puesta en marcha de este proyecto, se establecieron dos etapas:

Un primer momento en el que se identifica el problema o la necesidad que se quiere solucionar dentro del repositorio que en este caso se hace por medio el coordinador del repositorio; seguidamente, se hace una verificación de documentos contenidos del repositorio, dentro de ellos son, el manual de ingreso, las políticas y la plataforma DSpace. Después, se atribuye a la identificación de materias desde el listado alfabético del repositorio en el que en esta etapa se hace el tratamiento terminológico de las 200 materias que requieren un proceso de control de calidad.

Y un segundo momento, que es hacer la gestión de cada materia por medio de herramientas y uso de lenguajes controlado, en este caso, las LEMB digital, se ayuda con el tesauro de la UNESCO, de otras ciencias interdisciplinares, entre otras en el que se deje el registro en un archivo en Excel con la estructura del tesauro.

En segunda medida, establecer el control de autoridades para los autores en el que hace una identificación de los primeros 200 autores del repositorio y se hace un tratamiento con los nombres de los autores que requieren proceso de calidad, aquí se trata de que se deje una estructura en Excel, este proceso de validación de autores se complementa con herramientas como VIAF, catálogos públicos en la línea de bibliotecas y Cvlac.

Las materias y los autores se comprueban de acuerdo a lo que establece el Manual de Políticas para el ingreso de la información del proceso de gestión de información el cual, se desea que se corrijan los registros en DSpace, registrar los términos en el tesauro gráfico, usar herramientas de lenguaje controlado y validar tandas Conjunto de 20 Metadatos.
Finalmente, obtener un control de calidad dentro del repositorio.

5. Desarrollo de actividades

En esta etapa se presenta específicamente como fue el desarrollo de la metodología, en la que se logró el objetivo principal así:

Para la primera parte de la práctica en el Repositorio Institucional de la Universidad de Antioquia se abordó de manera general una contextualización de lo que son los repositorios, en lo que se trató básicamente de una contextualización por medio de investigaciones referentes a lo que son los lenguajes controlados y como empezar a aplicarse al repositorio de la Biblioteca Carlos Gaviria Díaz de la Universidad de Antioquia, lo cual mejoró el panorama a lo que se quiso lograr que es el control de autoridades y estandarización de materias y autores.

Se identificaron y se revisaron las políticas del repositorio institucional, el manual de ingreso y la plataforma DSPACE los cuales, ayudaron al análisis de la información para establecer la manera en el que se muestran los contenidos en los documentos del repositorio.

Además, se organizaron los registros a los que se les aplicó el control de autoridades, los cuales, en una primera etapa fueron las materias en las que se organizó un listado en orden alfabético, en las que se lograron separar en un archivo de Excel las materias que se trabajaron y además que las materias subrayadas en color rojo son a las que se le aplicó el tratamiento completo de la organización (*Ver Anexo 1. Listado de materias)* y seguidamente se avanzó en el trabajo de la construcción del tesauro gráfico del repositorio, en el que se empleó un tratamiento terminológico a 200 materias que requirieron de un proceso de calidad; este paso también se desarrolló en un formato de Excel en el que se hizo la gestión de información de cada materia y se aplicaron los lenguajes controlados el cual, se desarrolló con el uso de las herramientas, LEMB digital y otros tesauros específicos como el de la Unesco, y el tesauro del Centro de Documentación los cuales ayudaron a la verificación de los términos y se establecieron las siguientes convenciones:

TERI → Termino extraído del Repositorio Institucional

Término en inglés validado de las LEMB digital

U → Use

TE → Término específico

UP → Usado por

TG → Términos generales

F → Fuente de consulta y validación

NOTAS → Referentes al término o al registro encontrado

Número de registros → Cantidad de registro que fueron intervenidos.

De manera conjunta se hizo una construcción de atributos en los lenguajes documentales para la facilidad del acceso y la recuperación de la información en bases de datos y repositorios institucionales.

Dentro de las materias se hizo tratamiento a las que aparecen en los registros y, a su vez, se buscaban otros relacionados con esos registros o documentos a los que también se les aplicó el tratamiento, se trataba de que con los términos o materias que se relacionen, ponerlas en comillas para recuperar mucho más los registros de Dspace y hacer un control de autoridades de los registros mucho más completo con cada materia que aparezca.

Se organizaron los nombres en los metadatos de los documentos en la parte inferior inicialmente de las materias de los Anales de la Academia y los boletines de Medicina que son registros a los que se les cambió el metadatos de los nombres de los registros. Estos hacen parte de los artículos de revistas y patrimonio de la Universidad. Se anota que la organización de estos registros como quedan de manera correcta deben verse en los anexos.

Se organizaron de manera completa algunos metadatos como el resumen o el abstract en donde se encontraron errores mínimos de ortografía, títulos y nombres de los registros del repositorio. A algunos códigos *handle,* que es la URL de los registros del repositorio, se les hizo modificaciones y se nombraron alguno de ellos en la siguiente tabla:

http://bibliotecadigital.udea.edu.co/handle/10495/6307

Tabla 1. Códigos handle del repositorio

	3599	222	4076
4078	2465	370	3422
6333	362	5533	67
4317	4067	192	4065
3061	644	3593	4317
5655	2812	1504	3430
3497	1448	532	593
4705	7150	4112	4691
3584	1830	7000	4595
4694	4365	3604	4066
4396	4566	4153	3396
2298	4389	3629	4135
4394	3423	155	4569
3580	3408	4220	20
4015	3419	6933	1242
628	3517	3425	3155
4064	5015	2260	2289

Fuente: Elaboración propia

Luego se decidió que para darle más control de autoridades a los registros se debe establecer la relación con las terminologías que los involucran.

En la segunda etapa de la práctica académica se logró el objetivo que se establece, que es el de identificar los 200 autores del repositorio en los que se denotó la corrección de cada uno respecto a las fuentes de validación del autor como OPAC, SCOPUS, WORLDCAT, RESEARCHGATE, VIAF, CVLAC o GOOGLE ACADÉMICO y se aplicó el mismo ejercicio que se hizo con las materias, en el que los autores que se trabajaron se organizaron en un documento en Excel y se les aplicó una estructura de autores identificados del repositorio, fuentes de validación, registros asociados al autor, autor estandarizado y los números de registros que fueron intervenidos. (*Ver Anexo 9. Identificación de autores*)

Se tomó la estructura que establecen las normas de catalogación y como se establece en el manual de ingresos de contenidos de la Biblioteca Carlos Gaviria Díaz, además de los parámetros de búsquedas en las diferentes fuentes de validación especialmente para el CVLAC de Colciencias, se aplicó de la siguiente forma que es el nombre completo del autor entre comillas y se agrega la palabra CVLAC en el buscador de Google.

Así: “Nombre Apellido” Cvlac

Se estableció además un listado de autores los cuales se inició con la verificación por OPAC, SCOPUS, WORLDCAT, RESEARCHGATE, VIAF, CVLAC Colciencias y GOOGLE ACADÉMICO; debido a que ya aparecían unos que estaban organizados y corregidos y se denotaron con el color rojo, los que no aparecian de la verificación de ninguna de las fuentes de color verde, los que se validaron y autorizaron de color morado y finalmente los de color azul que fueron los que están validados y dentro del repositorio estaban escritos de manera correcta. *(Ver Anexo 10. Convenciones descritas para los autores)*

Se verificó un registro y en la búsqueda apareció que estaban dos (2) registros asociados a ese mismo autor y lo que se trataba es de copiar el mismo nombre con el que aparece en el registro de la búsqueda de las fuentes de validación para que aparezcan los registros completos del autor y se vean organizados.

Además a los autores que tenían un rol diferente dentro del documento se lograron recuperar de la misma manera que lo establecen las normas de catalogación y se organizaron según se describen en un metadatos de notas. *(Ver Anexo 11.)*

Se anotó la forma en como aparecían los autores en el repositorio antes de ser organizados así:

Rodríguez Correa, Ramón (Introducción/Prólogo)

a, e, i, o, u Producciones (Autor)

La Chiva (Periódico)

ADN (Periódico)

Q’ hubo (Periódico)

El Espectador (Periódico)

Y así como estos registros aparecían, había otros que también se les debe seguir haciendo una validación e irse reorganizando.

Después de describirse en el metadatos de notas de manera correcta queda de la siguiente manera *(Ver Anexo12)*:

Descriptor.note: Introducción y Prólogo de Ramón Rodríguez Correa.

Descriptor.note: Guión y producción de: a, e, i, o, u Producciones y Universidad de Antioquia

Descriptor.note: Artículos de periódicos contenidos en: ADN, La Chiva, Q'hubo y El Espectador.

Cabe anotar que también se logró la corrección de otros registros validados por un auxiliar de la biblioteca, en los que él, identificó 47 autores que van desde la letra P hasta la letra V, en dichos registros se identificaron que los que aparecen es porque están mal escritos o porque pertenecen al mismo autor, además anota la forma correcta en la que va a aparecer el registro y los títulos relacionados. Con esto, lo que se logró es un control de calidad así como se estableció con los 200 autores correspondientes a la práctica, la validación de los autores en las diferentes fuentes utilizadas y además la corrección de los registros en DSpace.

6. Resultados

Los resultados de este trabajo cuyo objetivo principal es establecer un control de autoridades y estandarización de metadatos en las materias y en los autores que se hayan identificado del Repositorio Institucional de la Universidad de Antioquia.

Es un objetivo que se logró en su totalidad, ya que las actividades que se plantearon para llevar a cabo este proyecto fueron procesos enteramente satisfactorios.

Para obtener este resultado, se hizo la revisión del manual de políticas el cual se estableció como un paso primordial para el análisis de la información y ayudar de manera clara a establecer los contenidos de los documentos del repositorio.

También, el avanzar en la construcción de los atributos en los lenguajes controlados de las materias, en las que se facilitó el acceso y recuperación de la información contenida en bases de datos y repositorios institucionales y además lograr unificar dichos descriptores para que esté mucha más organización y validarse con las diferentes herramientas o fuentes usadas.

Finalmente, en el control de autoridades para los autores se logró un ejercicio importante, ya que el uso de las fuentes de validación de autores nos ayudó mucho a constatar la escritura correcta de los mismos y lograr unificar los registros que pertenecen a un autor determinado.

A continuación, se mostrarán gráficamente los resultados que se obtuvieron después de haber realizado este trabajo dentro del repositorio:

Con las materias como se mencionó anteriormente en la metodología utilizada fue hacer un control autoridades a algunos registros del repositorio institucional, para ello, los datos que se obtienen es el de agregar términos al repositorio institucional donde lo que se hace es complementar cada registro y además estandarizarse con un lenguaje controlado.

La siguiente tabla, muestra cómo fueron los resultados que se obtuvieron de forma que explique gráficamente los registros que se lograron intervenir y de qué manera se obtuvo el resultado final de la práctica que es un total de 601 documentos los cuales se lograron intervenir de manera general.

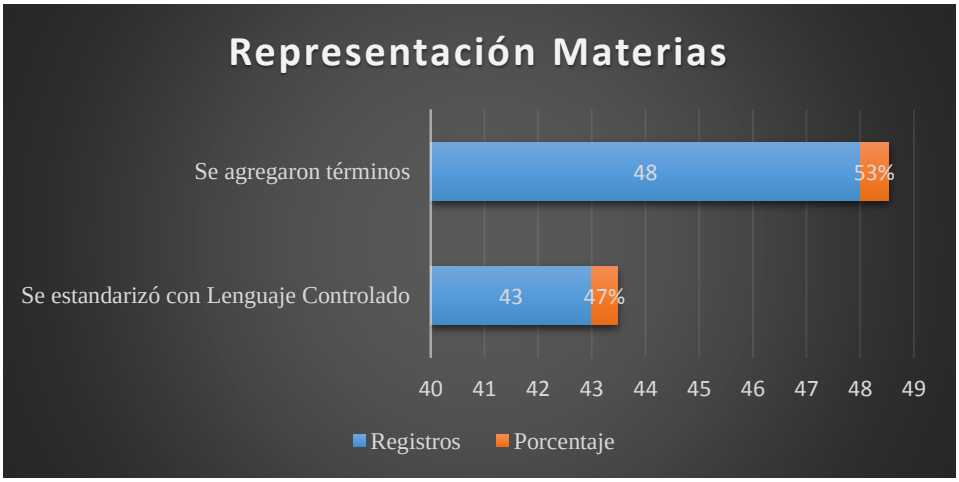
Tabla 2: Resultados control de autoridades

CONTROL DE AUTORIDADES	REGISTROS	PORCENTAJE	NRO. DOCUMENTOS ASOCIADOS A LOS TÉRMINOS
Se estandarizó con Lenguaje Controlado	43	47%	459
Se agregaron términos	48	53%	142
Totales	91	100%	601

Fuente: Elaboración propia

Es por ello, que la siguiente gráfica equivalente a la representación de las materias de forma cuantitativa muestra que en 48 registros analizados se agregaron términos y eso equivale al 53% de los términos que se agregaron al repositorio.
Seguidamente, muestra que 43 registros que se estandarizaron con un lenguaje controlado un 47% equivalen a los términos del repositorio institucional.

Gráfico 1. Representación de materias



Fuente: Elaboración propia

Además, se logró ver gráficamente que de un total de 459 números de documentos asociados a los términos 43 registros fueron estandarizados dentro del repositorio, es decir, que estos registros se cambión totalmente el término usado dentro del repositorio ya que la fuentes de

validación LEMB digital mostraba como era el término correcto o que se relacionaba con dicho documento.

Gráfico 2. Estandarización lenguaje controlado



Fuente: Elaboración propia

Para los autores, también se hizo lo mismo debido a que ayuda en gran medida a saber cuántos fueron los registros intervenidos y así poder cumplirse con el objetivo final.

Inicialmente, se elaboró una tabla donde se usan unas convenciones o colores para poder diferenciar los autores que están escritor de manera correcta en el repositorio, sin fuentes de validación, los que son validados y autorizados y finalmente los que se confirman que están correctos para un total de 200 autores intervenidos.

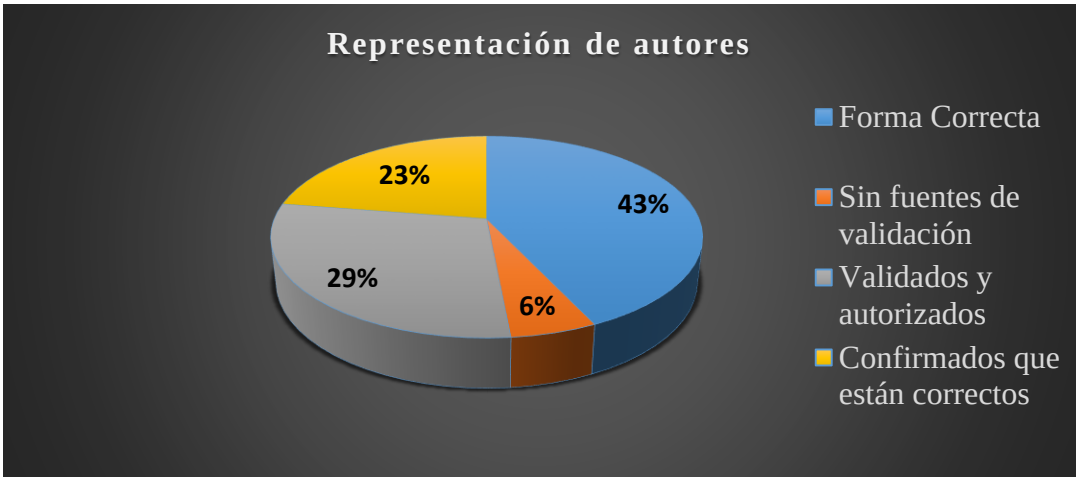
Tabla 3. Registros validados

Convención	Forma de la convención	Registros Validados	Porcentaje equivalente
	Forma Correcta	86	43%
	Sin fuentes de validación	11	6%
	Validados y autorizados	58	29%
	Confirmados que están correctos	45	23%
Total		200	100%

Fuente: Elaboración propia

Se denotó que autores están de forma correcta de 86 registros validados en el repositorio el 43% de los registros equivalen a los registros intervenidos, de 11 registros validados un 6% de los registros equivalen a los registros que están sin fuente de validación, es decir, que de las fuentes utilizadas ninguno dio información sobre el autor.
De 58 registros validados y autorizados le corresponde un 29% de los registros abordados y finalmente, 45 registros confirmados que están correctos, es decir, que están bien escritos en el repositorio equivalen a un 23%.

Gráfico 3. Representación de autores



Fuente: Elaboración propia

Para los autores, además de lograr como es la representación de manera gráfica, también se realizó un ejercicio parecido con las fuentes de validación utilizadas con los diferentes autores y así poder mostrar también cuales son las fuentes en las que se encontraron.

Se encontraron que 24 autores se validaron por Cvlac Colciencias que equivale a un 37%, 13 autores fueron validados por el OPAC de la Universidad de Antioquia que equivale al 20%, 4 autores por WorldCat que equivale al 6%, 10 autores validados en VIAF que corresponde al 15%, 2 autores validados en ResearchGate que le corresponde el 3%, otra fuente de validación 4 autores que equivalen a un 6%, 3 autores validados en Google Scholar que le corresponde el 5% y en Scopus 5 autores que equivalen al 8% para un total de 65 autores encontrados en estas fuentes.

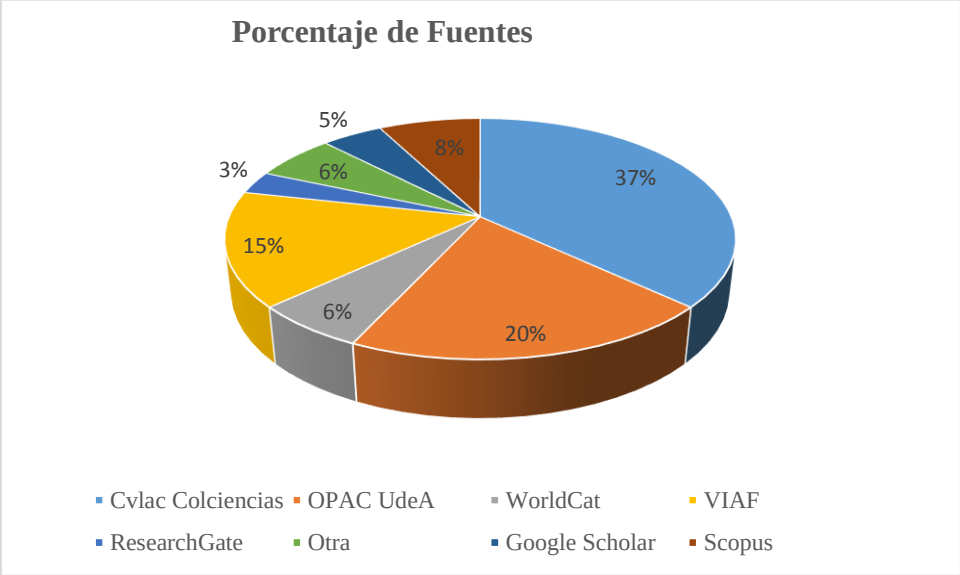
Tabla 4. Fuentes de validación

Fuentes de validación	***Cantidad***	**Porcentaje**
Cvlac Colciencias	*24*	37%
OPAC UdeA	*13*	20%
WorldCat	*4*	6%
VIAF	*10*	15%
ResearchGate	*2*	3%
Otra	*4*	6%
Google Scholar	*3*	5%
Scopus	*5*	8%
Total	*65*	100%

Fuente: Elaboración propia

En esta gráfica se muestra de manera gráfica el ejercicio con las fuentes de validación utilizadas.

Gráfico 4. Porcentaje de fuentes de validación



Fuente: Elaboración propia

7. Conclusiones y recomendaciones

El control de autoridades en un repositorio institucional es primordial para lograr que se normalicen los contenidos y a su vez, que haya relación con los diferentes puntos de accesos.

Es un trabajo que se hace con el fin de conocer más a fondo los repositorios, que no es solamente el ingresar la producción académica de una institución, sino que permite visibilizar el trabajo de la comunidad académica como tal; es por ello que el control de autoridades se constituye en una base vital para el repositorio.

Este trabajo ayudó a comprender que es algo que se debe seguir manejando en la institución debido a que son muchas las producciones de la universidad y aunque los centros de documentación ayudan un poco debido a que sus auxiliares ponen de su parte para poder ingresar los contenidos, el repositorio institucional es la cara de todo ese proceso, entonces se puede decir, que a modo de ver es recomendable que se les dé las pautas adecuadas a quienes ayudan en el repositorio de las diferentes facultades en cuanto a su personal tales como sus auxiliares o personas que colaboran con este proceso para que así se pueda lograr un trabajo conjunto y mucho más completo.

8. Referencias

Biblioteca Nacional de España. (23 de Agosto de 2010). *Presentación y uso de VIAF*. Obtenido de http://www.bne.es/export/sites/BNWEB1/es/LaBNE/Docs/UsodeVIAF.pdf

Cardona Quintero, W. d., & Zapata Casas, L. M. (2016). *Maual abreviado para orientar el ingreso de contenidos a la Biblioteca Digital de la Universidad de Antioquia.* Universidad de Antioquia, Sistemas de Bibliotecas, Medellín.

Colciencias. (s.f.). *Colciencias*. Obtenido de http://www.colciencias.gov.co/sites/default/files/upload/glosario-colciencias.pdf

ELSEVIER. (2018). *Scopus*. Recuperado el 28 de Febrero de 2018, de ELSEVIER: https://www.elsevier.com/americalatina/es/scopus

Estrada Villacís, M. E. (30 de Diciembre de 2015). *Infotecarios*. Obtenido de http://www.infotecarios.com/lenguajes-documentales-y-lenguaje-natural-en-la-recuperacion-de-informacion/

Herrero Pascual, C. (1999). El control de autoridades. *Anales de Documentación*, 121-136. Obtenido de http://www.redalyc.org/html/635/63500208/

Juncà Campdepadrós, M. (2010). *Listas de encabezamientos de materia y listas de autoridades.* Obtenido de https://goo.gl/3cyxeZ

Moya Orozco, V. (9 de Febrero de 2016). *ResearchGate: la red social de los investigadores*. Obtenido de https://www.researchgate.net/publication/293654006_ResearchGate_la_red_social_de_los_investigadores_v_10#pf3

Muñoz de Frutos, A. (24 de Octubre de 2017). *Computer hoy*. Obtenido de https://computerhoy.com/noticias/internet/que-es-google-academico-como-funciona-52486

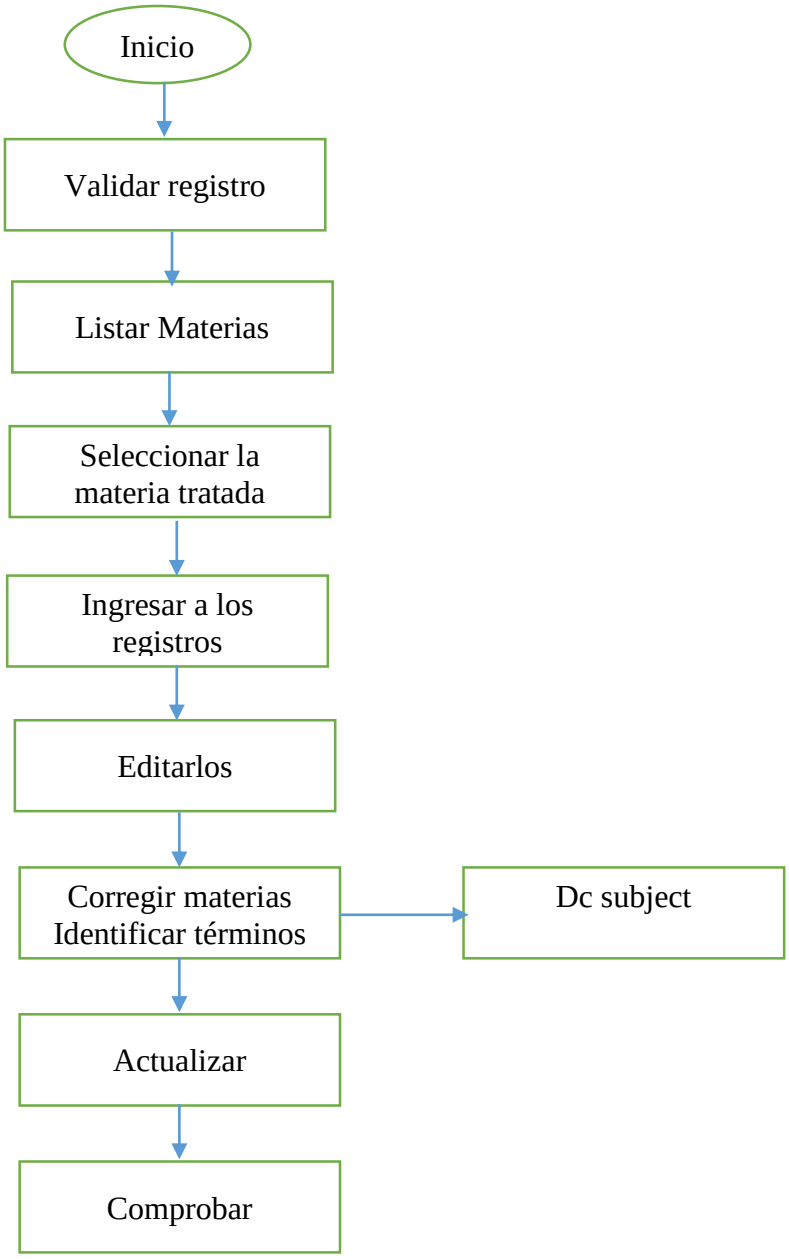
Nieto, S. y. (2013). *Control de autoridades en Dspace. en III.* Conferencia de Bibliotecas y Repositorios Digitales de América Latina (BIREDIAL '13), Universidad de Costa Rica. Obtenido de http://eprints.rclis.org/20560/1/Ponencia%2021-277-1-PB.pdf

OCLC. (2018). *Woldcat*. Obtenido de https://www.oclc.org/es/worldcat.html

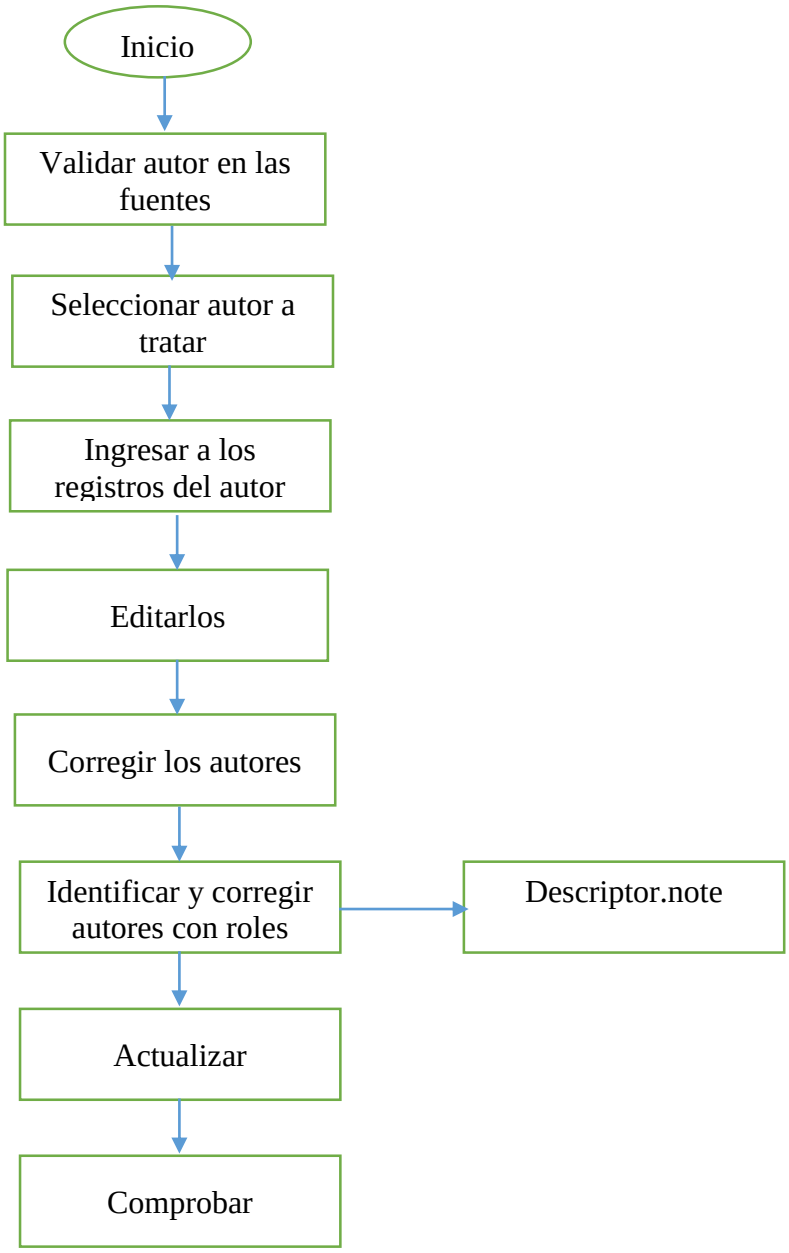
Sistema de Bibliotecas. (2017). *Catálogo Público OPAC*. Obtenido de Universidad de Antioquia: https://goo.gl/9Vgrhx

Sistema de Bibliotecas. (2017). *Manual de políticas para la organización y tratamiento de la información en la base de datos OLIB.* Sistemas de bibliotecas, Vicerrectoria de Docencia. Medellín: Sistema de Bibliotecas.

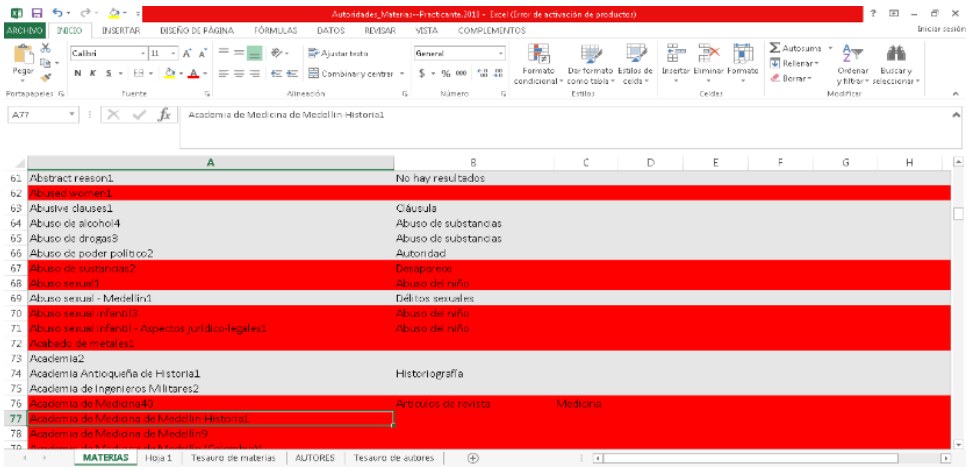
Anexo 1. Validación de materias



Anexo 2. Validación de autores



Anexo 3. Listado de materias



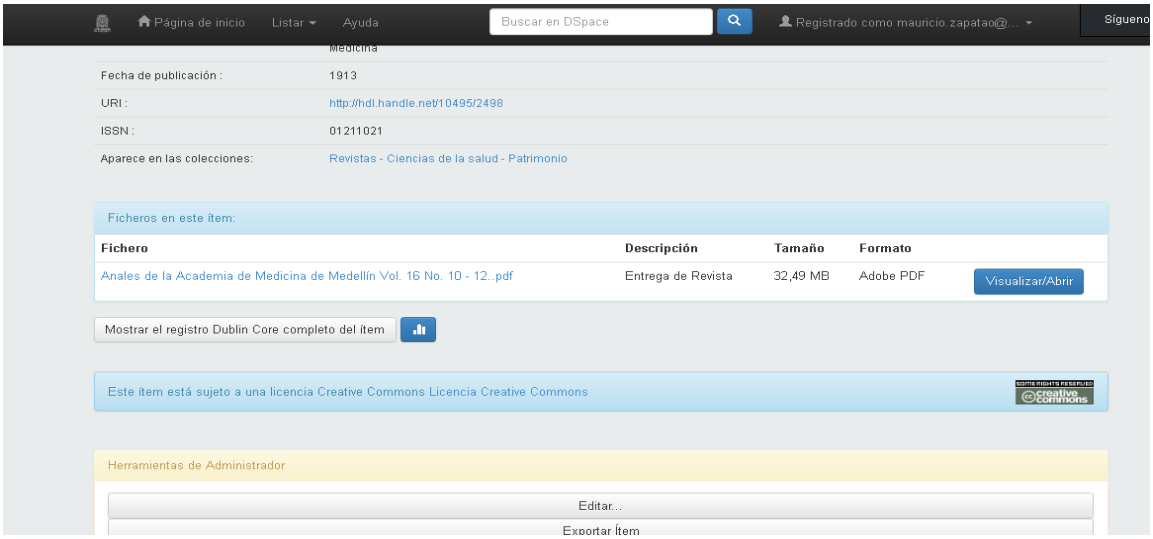
	A	B
61	Abstract reason1	No hay resultados
62	Abused women1	
63	Abusive clauses1	Cláusula
64	Abuso de alcohol4	Abuso de substancias
65	Abuso de drogas3	Abuso de substancias
66	Abuso de poder político2	Autoridad
67	Abuso de sustancias2	Desaparecer
68	Abuso sexual1	Abuso del niño
69	Abuso sexual - Medellín1	Delitos sexuales
70	Abuso sexual infantil3	Abuso del niño
71	Abuso sexual infantil - Aspectos jurídico-legales1	Abuso del niño
72	Acabado de metales1	
73	Academia2	
74	Academia Antioqueña de Historia1	Historiografía
75	Academia de Ingenieros Militares2	
76	Academia de Medicina40	Artículos de revista Medicina
77	Academia de Medicina de Medellín Historia1	
78	Academia de Medicina de Medellín9	

Anexo 4. Tesauro de materias construido

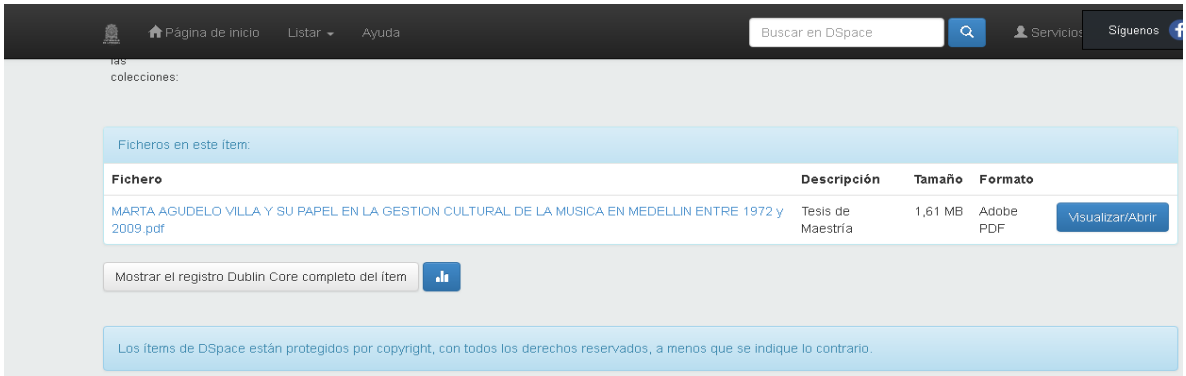
Termino del Repositorio Institucional	Use	Correcto	Término en inglés	Término Específico	Usado por	Término General	Fuente de Consulta y Validació	NOTAS	Número de Registros Intervenidos
TERI	U	C	TI	TE	UP	TG	F		
11 de Septiembre de 2001	Septiembre 11 (Ataque terrorista, 2001)	Septiembre 11 (Ataque terrorista, 2001)	September 11 Terrorist A	N/A	*Pentágono-Ataque terrorista, 2001	*Secuestro de aviones - Estados Unidos Hijacking of aircraft - United States *Terrorismo - Estados Unidos	Lemb		1
Abandono infantil	Abandono del hogar	Abandono del hogar	N/A	N/A	N/A	N/A	N/A	Abandono del hogar	3
N/A	N/A	Abandono del hogar	* Desertion and non-support	* Patria potestad * Obligación alimentaria	* Familia – Abandono Hogar – Abandono	* Derecho de familia Family law * Divorcio * Divorcio-legislación * Relaciones familiares	Lemb		3
Abandono de los estudiantes	Deserción universitaria	N/A	N/A	N/A	N/A	N/A	N/A		
N/A	N/A	Deserción universitaria	College dropouts	N/A	N/A	* Asistencia a la universidad College attendance * Deserción escolar *Orientación estudiantil en educación superior	Lemb		6
					*Niños-Aislamiento s	*Trabajo social con niños Child welfare Social work with children	Lemb		

MATERIAS | Hoja 1 | Tesauro de materias | AUTORES | Tesauro de autores

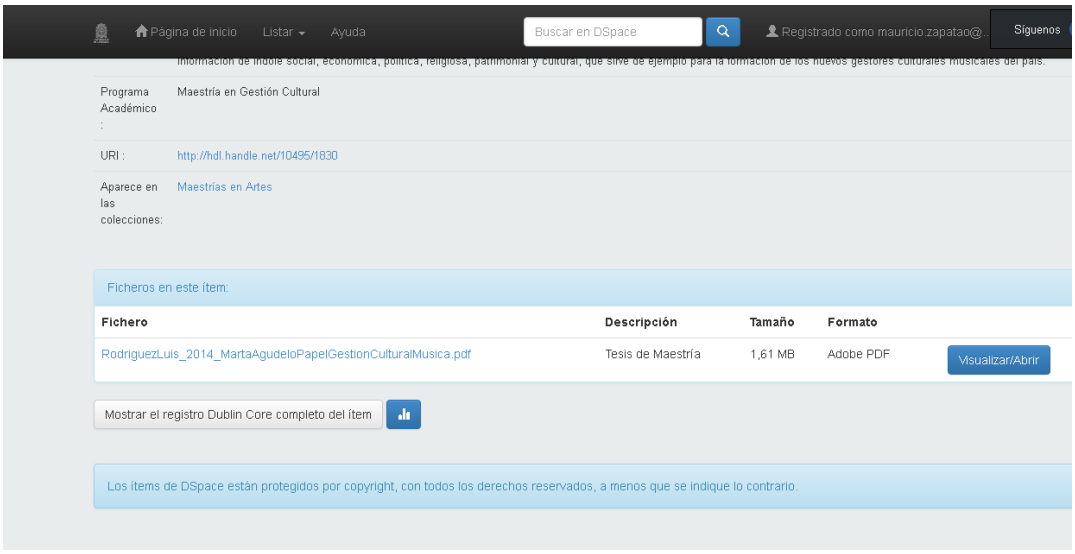
Anexo 5. Nombres que se logran organizar



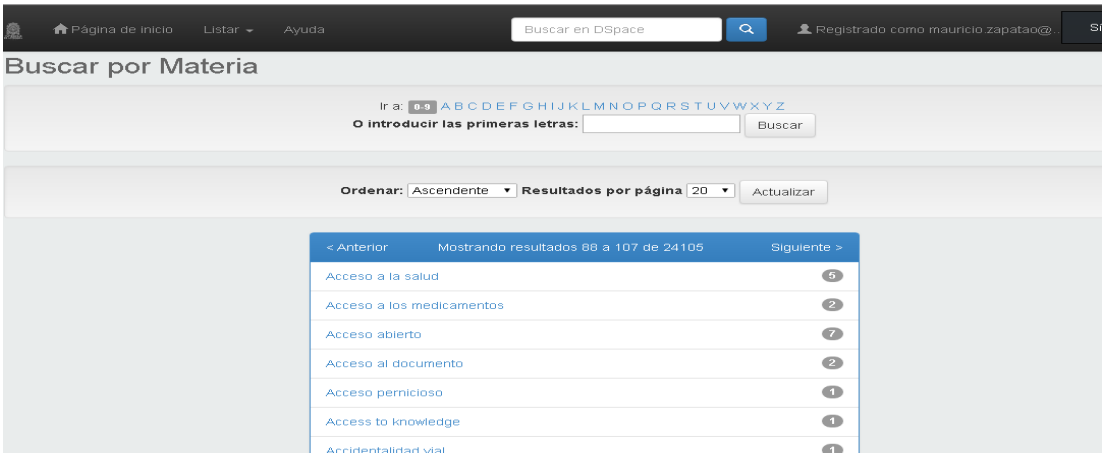
Anexo 6. Otros nombres



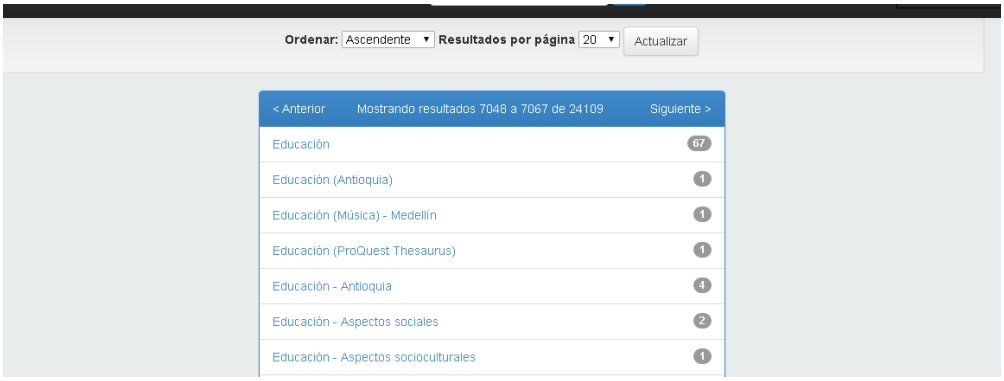
Anexo 7. Nombres organizados



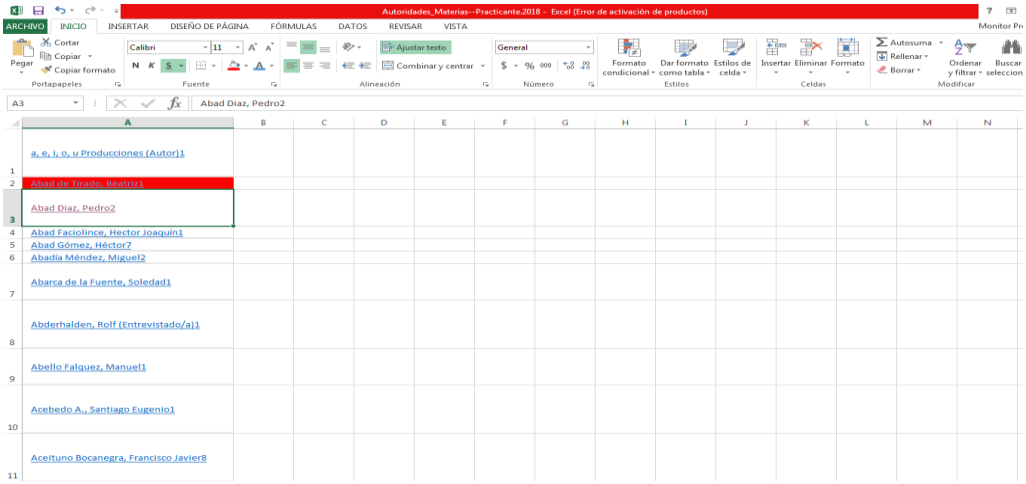
Anexo 8. Términos relacionados por materias unificados



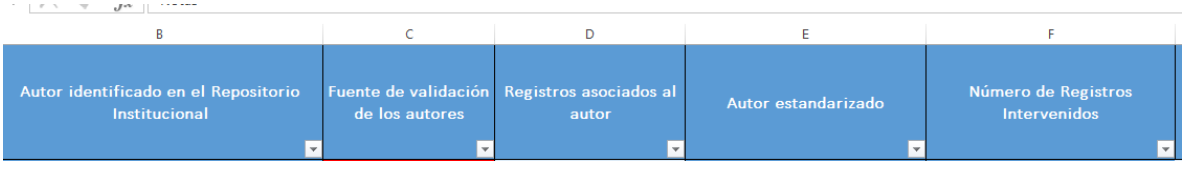
Anexo 9. Otros términos relacionados por materias



Anexo 10. Autores



Anexo 11. Identificación de los autores



Anexo 12. Convenciones descritas para autores

Convención	Forma de la convención
	Forma Correcta
	Sin fuentes de validación
	Validados y autorizados
	Confirmados que están correctos

Anexo 13. Autores descritos con el rol en los registros

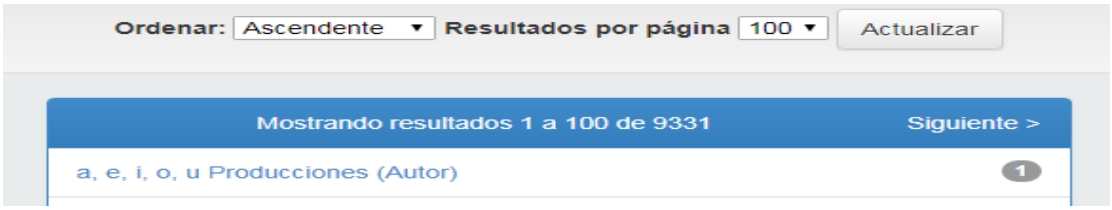
Por favor, use este identificador para citar o enlazar este ítem: http://hdl.handle.net/10495/3276

Título :	Cosas que fueron: colección de artículos de literatura, costumbres, critica y viajes
Autor :	Alarcon, Pedro Antonio de
Asesor :	Rodríguez Correa, Ramón (Introducción/Prólogo)

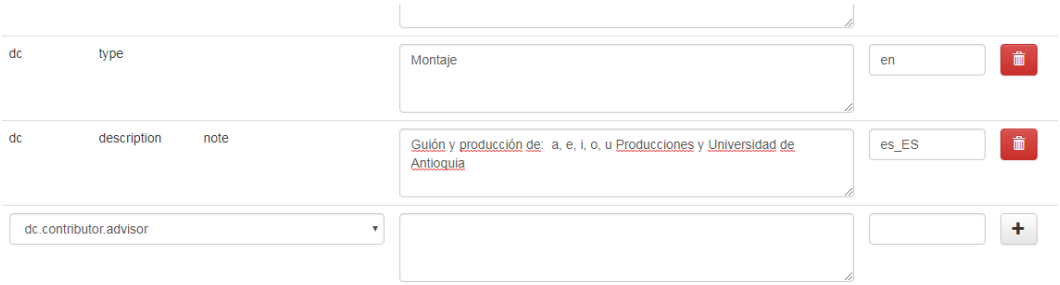
Anexo 14. Autor organizado en metadatos de notas

dc	date	issued	1871
dc	description	note	Introducción y Prólogo de Ramón Rodríguez Correa
dc	description	provenance	Submitted by Montoya Tamayo Consuelo Edilma (consuelo.montoya@udea.edu.co) on 2015-10-20T20:14:48Z No. of bitstreams: 2
dc	description	provenance	Approved for entry into archive by Weimar Cardona (weimard.cardona@udea.edu.co) on 2015-10-27T22:53:30Z (GMT) No. of bitstreams: 2

Anexo 15. Otros autores descritos con los roles



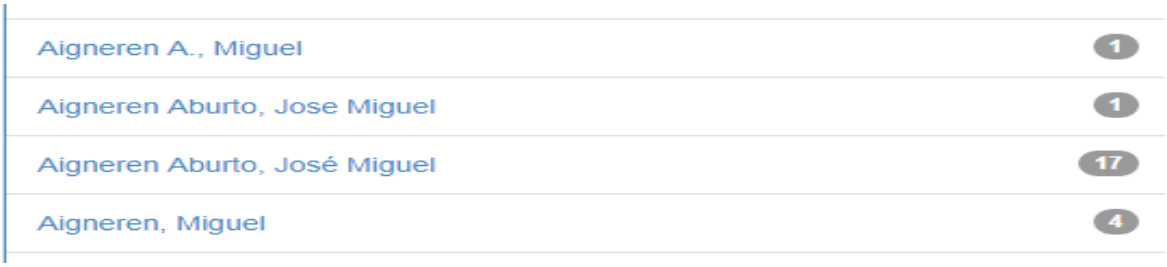
Anexo 16. Autores organizados en metadatos de notas



Anexo 17. Otros autores con roles descritos



Anexo 18. Varios nombres de autores



Anexo 19. Nombres de autores unificados

