



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

**POSGRADO EN BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN
FACULTAD DE FILOSOFÍA Y LETRAS
INSTITUTO DE INVESTIGACIONES BIBLIOTECOLÓGICAS Y ESTUDIOS DE LA
INFORMACIÓN**

Repositorios institucionales: un acercamiento a su análisis global

TESIS

QUE PARA OPTAR POR EL GRADO DE:

**MAESTRA EN BIBLIOTECOLOGÍA
PRESENTA:**

Cecilia Aurora Aguilar Rocha

**ASESOR: DR. Filiberto Felipe Martínez Arellano
Instituto de Investigaciones Bibliotecológicas y Estudios de la Información**

México, D.F. Marzo, 2015



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Tabla de contenido

	Pág.
Introducción	
Capítulo 1 Acceso abierto	1
1.1 Antecedentes del acceso abierto	1
1.2 Surgimiento del acceso abierto	8
1.2.1 Iniciativas en favor del acceso abierto	10
1.2.1.1 Declaración de Budapest	11
1.2.1.2 Declaración de Bethesda	13
1.2.1.3 Declaración de Berlín	15
1.3 Concepto	18
1.4 Ventajas del acceso abierto.....	24
1.5 Situación actual	29
Capítulo 2 Repositorios Institucionales	33
2.1 Antecedentes de los repositorios institucionales	33
2.2 Características	36
2.3 Definiciones	40
2.4 Creación de repositorios institucionales	43

	Pág.
Capítulo 3 Repositorios institucionales: análisis de la situación internacional	49
3.1 Introducción	49
3.2 Materiales y métodos	51
3.3 Resultados	55
3.4 América	56
3.5 Europa	57
3.6 Asia	58
3.7 África	59
3.8 Oceanía	59
3.9 América Central y El Caribe	60
3.10 América del Sur	60
3.11 Discusión	61
3.12 Proporción de repositorios por continente	66
3.13 Proporción de repositorios por países	68
3.14 Proporción de software utilizado en repositorios	69
3.15 Tipología en los repositorios institucionales	71
3.16 Conclusión	72

	Pág.
Capítulo 4	
Metadatos y software para repositorios institucionales	76
4.1 Metadatos	79
4.1.1 Definiciones y características	82
4.1.2 Tipos de metadatos	85
4.1.3 Aplicación de los metadatos	88
4.1.4 Esquemas de metadatos	91
4.2 The Dublin Core	94
4.3 Software	99
4.3.1 DSpace	101
4.3.2 EPrints	105
4.3.3 Greenstone	107
4.4 Conclusión	110
Conclusiones	116
Obras consultadas	122

Se presenta un análisis general de los repositorios institucionales implementados a nivel mundial, tomando como base la información que presenta Opendoar, así también se analizan los elementos más utilizados que forman parte de la infraestructura para su óptimo funcionamiento.

Los repositorios generalmente se han implementado en instituciones académicas y es factible observar que sus contenidos consisten en grandes volúmenes de documentos científicos en diferentes formatos. Los resultados obtenidos en este estudio manifiestan que los países desarrollados son líderes en la implementación de repositorios digitales, el software más usado para el funcionamiento de estas herramienta es DSpace y además es posible observar que los repositorios institucionales contienen información puramente académica.

Palabras lave: acceso abierto, repositorios institucionales, software, metadatos

A general analysis of repositories implemented globally, based on the information presented OpenDOAR and frequently used items that are part of the infrastructure for optimum performance are also analyzed is presented.

The repositories are usually implemented in academia and is possible to observe that its contents consist of large volumes of scientific documents in different formats. The results obtained in this study demonstrate that developed countries are leaders in the implementation of digital repositories, the most used software for running these tools is DSpace and it is possible to observe that institutional repositories contain purely academic information.

Words wash: open access institutional repositories, software, metadata

Introducción

En la era de las redes y tecnologías de la información, favorecer la comunicación más abierta posible a los resultados de la investigación científica, ha traído compromisos por parte de los científicos, instituciones y gestores de la información; de igual manera, ha llevado a profundas reflexiones en diversos ámbitos como son los relativos a la calidad, visibilidad e impacto de los resultados de investigación, la gestión de los derechos de autor y la garantía de preservación digital.

Las Tecnologías de Información Comunicación (TIC), han transformado la manera en que se procesa, se guarda, se accede, se comparte y se analiza la información; así mismo han introducido cambios de importante alcance, trayendo como factor fundamental el cambio de soporte, es decir, el cambio del formato tradicional impreso al digital.

El Acceso Abierto ha sido aceptado ampliamente por muchos interesados en obtener información académica y ha llevado a una renovación de los modelos de comunicación científica tradicionales y el planteamiento de alternativas a los modelos ya existentes.

Los repositorios institucionales tienen gran importancia para la comunicación científica; en este sentido, es una alternativa para aprovechar las posibilidades que ofrece Internet en la difusión del conocimiento, más allá de las restricciones marcadas por los intereses comerciales. Estos repositorios -creados y mantenidos por la universidad que autoriza su implementación-, recogen los contenidos generados por los miembros de su comunidad y para su funcionamiento se han desarrollado diversas aplicaciones destacándose el software que puede ser de libre acceso o comercial.

Otro aspecto necesario para el funcionamiento de los repositorios institucionales que garanticen la interoperabilidad, es el uso de metadatos basados en estándares internacionalmente normalizados y el software libre del que hacen uso.

Por su parte los repositorios institucionales se están convirtiendo progresivamente en lugares en los que numerosas instituciones almacenan y organizan el resultado de sus actividades. Se pueden entender como una alternativa o complemento al sistema tradicional de comunicación científica, que pretende recuperar la publicación de trabajos que faciliten su visibilidad y crítica a los autores por parte de sus colegas, con la intención de investigación, progreso y difusión del conocimiento. Por lo tanto, es factible reconocer que el aumento de repositorios institucionales se está dando continuamente, al grado de contar actualmente con más de dos mil de ellos (*OpenDOAR* (Directorio de Repositorios de acceso abierto) (Directory of open Access repositories (<http://www.opendoar.org/>)).

Para la implementación y desarrollo de los repositorios digitales es necesario contar con cierta infraestructura que garantice su óptimo funcionamiento, por ello, el propósito de este trabajo es presentar un análisis general de los repositorios institucionales implementados a nivel mundial, tomando como base la información que presenta *OpenDOAR*. Así también analizar los elementos más utilizados que forman parte de la infraestructura de los repositorios para su óptimo funcionamiento.

Por otro lado, es factible observar que sus los contenidos de los repositorios consisten en grandes volúmenes de documentos académicos en diferentes formatos, por eso también se analizan los contenidos que forman parte de estos repositorios a nivel mundial, con el propósito de comprobar que los repositorios son usados académicamente.

Lo anterior permitirá continuar la realización de otros estudios relacionados con los repositorios institucionales de forma más específica como: la definición de políticas

de uso, políticas de calidad, políticas de preservación digital, políticas de normalización de metadatos, políticas de propiedad intelectual, entre otras. Ésto impulsará la implementación de repositorios académicos, cuya utilidad redundará en el uso continuo de ellos. Para eso será necesario darlos a conocer; de esta manera, habrá la necesidad de capacitar tanto a los investigadores como a los profesionales de la información en temas teóricos y prácticos relacionados con el modelo de acceso abierto.

El estudio se divide en cuatro capítulos orientados al logro de los propósitos antes mencionados.

Este trabajo comienza (capítulo 1) con una breve introducción histórica que permite conocer los conceptos relacionados con el Acceso Abierto, las iniciativas que a él conciernen y las ventajas de este movimiento en favor de tener la información científica en libre acceso. Además se da un breve comentario referente a la situación actual que guarda el Acceso Abierto originado por inconformidades derivadas de los científicos hacia diferentes editoriales comerciales.

En el siguiente capítulo (2), se exponen algunos aspectos generales de los repositorios institucionales, sus antecedentes, características, su implementación además de algunas reflexiones generales relacionadas con el software recomendado.

En el capítulo 3, se describen los resultados obtenidos del análisis basado en el directorio OpenDOAR (The Directory of Open Access Repositories-OpenDOAR) (Directorio de repositorios de acceso abierto) en línea. Este análisis se realizó con la intención de conocer la situación que contemplan los diferentes repositorios institucionales a nivel internacional. Además se muestran gráficamente los resultados obtenidos del análisis.

Por último, en el capítulo 4 se indican algunos conceptos generales de los metadatos, un análisis concerniente a los esquemas de metadatos y su relación con las plataformas de software más utilizados en la implementación de repositorios institucionales.

Finalmente, se exponen las conclusiones derivadas de la investigación.

Capítulo 1

Acceso Abierto

Los interesados en obtener información relacionada con los avances científicos y humanísticos, pensaron en las posibilidades que ofrece la tecnología para difundir sus resultados intelectuales de la ciencia y favorecer la investigación, abriendo la posibilidad de que el conocimiento pudieran estar disponible al resto de la comunidad científica de manera universal e inmediata, pero los resultados no fueron los esperados, pues continuaron las barreras, sobre todo económicas, a causa del abusivo aumento de los precios relacionados con las suscripciones de las revistas que interpusieron las grandes empresas editoriales. Ante la necesidad de obtener información actualizada y con mayor facilidad, se desencadenaron numerosas manifestaciones en favor del movimiento por el acceso abierto dando figura a tres declaraciones fundamentales, las cuales han tenido más influencia en la configuración y definición de este movimiento. Estas declaraciones son: Declaración de Budapest, Declaración de Bethesda y la Declaración de Berlín.

1.1 Antecedentes del acceso abierto

Las actitudes y conductas varían de una disciplina a otra así pues, en las áreas científicas la emergencia por la necesidad de comunicar en tiempo reducido resultados recientes, forma parte de ser de la investigación. Es así que los científicos relacionados con las ciencias y técnicas han estado más a favor del acceso abierto por la necesidad de la investigación que generan, -tomando en cuenta que necesitan información actualizada para el desarrollo y cumplimiento de sus proyectos en favor de la humanidad-. En cambio, el problema de conseguir información actualizada ha sido en menor medida para los involucrados en

ciencias sociales y humanísticas, si bien, aunque sus investigaciones puedan basarse en libros y no necesariamente en las últimas investigaciones, en ocasiones es necesario hablar de acontecimientos actuales contenidos en documentos difíciles de conseguir.

El acceso abierto (OA, por su forma en inglés) es una respuesta de las comunidades científicas a los grandes monopolios de empresas comerciales, quienes concentran las más prestigiosas y esenciales publicaciones que ayudan a la difusión de nuevas ideas. Es una manifestación en contra del aumento de los altos precios a las revistas científicas, mismos que han dificultado el acercamiento a los resultados de investigación y sus productos.

El acceso abierto tiene como objetivo mejorar la comunicación científica y eliminar barreras que impidan el acercamiento a la información, maximizando la generación de recursos electrónicos de acceso abierto sin costo y como propósito tiene: dar una alternativa viable que, aprovechando al máximo las tecnologías de la información y la comunicación, amplía la capacidad de distribución y reduce los costos con el fin de proporcionar un más amplio y fácil acceso a los resultados de la investigación limitando las restricciones impuestas por los editores comerciales.

El acceso abierto es una alternativa que por muchos años ha provocado opiniones por parte de investigadores y académicos, en virtud de que generalmente las publicaciones científicas se han manejado como un negocio rentable para las editoriales, quienes han puesto barreras para hacer la información accesible a los interesados.

Parte del siglo XIX y la totalidad del XX han traído como resultado diversos factores que condicionaron lo que se denominó **crisis de la comunicación académica** o **crisis de las revistas tradicionales** en la información científica impresa (Chan, 2004, p. 278); entre esos factores, Sánchez Tarragó (2009), señala como más importante el incremento sostenido de los precios de las revistas

científicas, sobre todo en las áreas de ciencia, tecnología y medicina, y la necesidad de comunicar los resultados de forma inmediata.

Con la llegada de la Web se dio el hecho de que la información se distribuiría a los doctos en todas las áreas del conocimiento, suponiendo la eliminación de los costes de impresión y suscripciones; sin embargo, la alternativa elegida por los editores era pagar por suscripción en línea tanto o más que por la edición impresa.

Con el tiempo las ganancias obtenidas de las publicaciones académicas han aumentado en beneficio de las casas editoriales, no importando que las aportaciones intelectuales para la publicación de las revistas sean de los investigadores, de tal manera que esos resultados intelectuales terminan siendo comercializados por esas editoriales que obtienen amplios beneficios por su edición. Esto es sabido por los intelectuales quienes adquieren un compromiso al desempeñarse en una institución, pues por sus publicaciones reciben reconocimiento intelectual por parte de la comunidad científica con el ánimo de obtener impacto por la difusión de sus investigaciones y, como resultado, tener privilegios como son: presupuesto, apoyo de otras instituciones, algún tipo de promoción o mayor facilidad para conseguir financiación para investigaciones que se encuentren desarrollando, entre otros.

En virtud de lo anterior, los conocimientos científicos, técnicos y humanísticos, así como el principal medio de divulgación para la ciencia conocido como publicación científica y principalmente las publicaciones periódicas, no podían estar a merced de políticas ni monopolios editoriales que impidieran el acceso al conocimiento. Este acaparamiento comercial provocó que a nivel internacional los involucrados en la ciencia reaccionaran, dando lugar a distintas iniciativas en favor del acceso abierto, a causa de la política inflacionaria de las revistas científicas tradicionales tendiente a la constante subida de precios y a la necesidad, por parte de las instituciones, de conservar, preservar y poner a disposición de la comunidad investigadora el patrimonio intelectual.

Entre otros factores que motivaron la llamada crisis de las revistas se encuentran: los costos que fueron un factor desencadenante, la gran demora entre la fecha en que se escribe un artículo y éste aparece publicado en la revista y la política de embargo de muchos editores científicos que, a la fecha, supone una barrera para el libre y rápido acceso a la información. El embargo consiste en hacer los contenidos de las revistas públicamente accesibles, después de un período de tiempo establecido a partir de su edición, normalmente el período oscila entre seis meses y un año; por ende, lo antes mencionado trajo como consecuencia, el que los científicos buscaran nuevas alternativas de comunicación científica.

Otros aspectos relacionados con la crisis del sistema de comunicación científica, fueron las restricciones que establecen las legislaciones sobre el derecho de autor, el acceso y la diseminación de la información producida por los investigadores, mismos que han desvirtuado el objetivo primario referente a favorecer la comunicación científica y los propios derechos, que han estado más enfocados a la publicación en revistas “de impacto” que a la diseminación de los resultados en cualquier área del conocimiento expuestos libremente (Sánchez Tarragó, 2009, p. 2).

La literatura en acceso abierto dispuesta en Internet no es gratuita, es una modalidad en la que se eliminan las barreras económicas, y no es totalmente libre, pues cuenta con derechos reservados sobre los trabajos.

En la literatura disponible en Internet de acceso abierto, los derechos de explotación son los que condicionan la reutilización de los trabajos y su autoarchivo en repositorios digitales de acceso abierto. El derecho de autor otorga protección al autor para que goce de prerrogativas y privilegios de carácter personal y patrimonial, de este modo, divide dos tipos de derechos: derechos morales y derechos patrimoniales.

Los derechos morales, son derechos inalienables, imprescindibles, irrevocables e inembargables. Estos derechos otorgan el poder a los autores para determinar si su obra será divulgada y en qué formato; además tienen el derecho de oponerse a cualquier modificación. Lo que se dispone este artículo quiere decir que en ningún caso se puede transferir una obra a un tercero si el autor no lo permite, puesto que él es el único titular de su obra. Éste es el derecho que tiene todo autor a ser reconocido como tal y el derecho a la integridad de su obra. En este caso el autor produce su obra y la tiene bajo su custodia.

http://www.indautor.gob.mx/documentos_normas/leyfederal.pdf .

Los derechos patrimoniales son aquellos en los que el autor tiene derecho a que su obra sea reproducida, distribuida y comunicada públicamente. El autor puede explotar su obra o bien autorizar o prohibir su explotación, no dejando con esto de ser el titular de los derechos. Así, un autor puede transmitir con libertad sus derechos patrimoniales: trasladarlos o adjudicar licencias con exclusividad y no exclusividad de uso, durante un tiempo determinado y de manera onerosa, quedando determinados los montos, el procedimiento y los términos para el pago de remuneraciones. En dado caso el autor cede sus derechos a una editorial para que su trabajo se publique bajo ella.

http://www.indautor.gob.mx/documentos_normas/leyfederal.pdf

Cabe recordar que los derechos de autor constituyen uno de los derechos de propiedad intelectual que se crearon con el propósito de proteger al titular otorgándole derechos morales de una obra contra acciones indebidas que terceros podrían hacer a ella. Empero, en el caso de las revistas científicas, los derechos patrimoniales de los autores pasan a manos de los editores, por lo que el autor pierde el control sobre el uso posterior de su trabajo publicado pero no pierde los derechos morales. De esta manera, las transformaciones del entorno digital han abierto las posibilidades para que los autores distribuyan sus trabajos sin depender de editores comerciales en los cuales las restricciones que se imponen sobre los derechos patrimoniales aumentan.

Lo anterior, no significa que los intelectuales se nieguen a que la institución en la que desarrollan sus investigaciones, difunda los resultados obtenidos en alguna publicación, esta sea comercial o de acceso abierto. En el caso comercial, el investigador otorga los derechos a la universidad o centro en el que labora para que se ocupe de difundir esos resultados a editores comerciales que más convengan. En cambio en el conocimiento de libre acceso no se desconoce la autoría de los creadores y reconoce que muchas veces los fondos que facilitan la creación de ese intelecto son resultado de fondos públicos, por tanto, es justo retribuir a la sociedad las posibilidades de acercarse al conocimiento para fines no lucrativos. Ésto es, el autor otorga sus derechos patrimoniales al lugar en el que investiga, pues a él le sigue pagando su centro laboral, él tiene la obligación de construir el conocimiento y difundirlo.

Odlyzko, entre otros autores coinciden en que la crisis existente desde hace años en la comunicación científica y en las publicaciones especializadas, es causa de que no se han cumplido los objetivos primarios, en otras palabras: favorecer la diseminación y el intercambio de los resultados científicos para lograr la fertilización de la ciencia y del progreso científico, técnico y social de la humanidad (Chan, 2004, p. 277, Odlyzko, 2009, p. 32).

Resulta necesario tener en mente que no todos los editores son entidades comerciales con ánimo de lucro, teniendo en cuenta que una buena parte de ellos son institucionales, pues forman parte de universidades, sociedades científicas y asociaciones profesionales que no obtienen beneficios directos, sino que los ingresos que perciben por suscripción sirven para financiar y sostener las actividades de su organización.

Resulta también necesario conocer que en el modelo tradicional de edición científica, las comunidades de científicos están pagando al editor por el servicio de publicación y difusión a través de un modo indirecto que es la suscripción, en cambio, en el modelo de acceso abierto se paga por artículo publicado como

modelo de precio transparente, en el supuesto de que proporciona acceso universal, y el total del costo por artículo publicado es más bajo que en el modelo de suscripción.

Con la aparición de Internet se pensó que este instrumento permitiría la accesibilidad al conocimiento y abaratar su costo, pero hubo editoriales que pusieron sólo algunos títulos en línea permitiendo su consulta en texto completo sin ningún cargo, por lo que se pensó en la gratuidad, pero se observó que las revistas impresas y electrónicas se combinaban para obtener ganancia y cobrar un cargo adicional por la suscripción en línea, tanto o más que por la edición impresa, justificando que se proporcionaban otros servicios de valor añadido.

Uno de los acontecimientos que marcó la revolución de la información, al igual que la publicación de artículos de importancia científica y la comunicación entre interesados para obtener resultados confiables –con el propósito de comunicar las diferentes investigaciones por parte de los científicos- fue la aparición de la World Wide Web (WWW); sin embargo, este adelanto no representó avances para la comunicación científica, puesto que los editores cobraban por la publicación tradicional y la electrónica –revistas tradicionales en papel, convertidas a electrónica y comercializadas en ambos formatos-.

De igual manera, con las Tecnologías de Información y Comunicación (TIC), se logró facilitar la comunicación entre científicos o investigadores –tomando en cuenta que pertenecen a cualquier área del conocimiento (humanístico o científico)- gracias al uso de medios electrónicos que ayudaron a conseguir, con mayor facilidad, los textos científicos, no importando el lugar geográfico en el que se encontraran; así, aunque los avances científicos, tecnológicos y humanísticos estuvieron al alcance de los interesados que necesitaban información, se dejó brillar la existencia de diversos obstáculos como: conseguir publicaciones con recientes resultados por parte de editoriales comerciales y de prestigio, quienes

impedían a los interesados obtener información actual que sólo la facilitaban a cambio de comercialarla.

Dadas las condiciones que anteceden, un problema que encontraron los investigadores teniendo a Internet como herramienta, fue que a pesar de que los textos estaban en formato electrónico, no se podían consultar libremente, por el hecho de ser manejados por editoriales comerciales. Otro inconveniente que se presentó fue el que los textos eran requeridos en formatos digitales para consultarlos en línea, lo que provocó la necesidad fundamental en los equipos de cómputo referente a su actualización, además de proveerlos para su interconectividad. También fue necesario poner al día los formatos y estándares en los que se almacenaba la información.

En los años 80 y 90 la mayor parte de las editoriales convirtieron en electrónicas sus revistas editadas en papel, comercializándolas en ambos formatos y fue así que la comunidad bibliotecaria se percató de que los editores no abandonaban sus publicaciones en papel por las electrónicas, sino que los conjugaban obligando a pagar un cargo adicional por la versión en línea. Para tratar de resolver el problema de adquisición del material, un grupo de científicos, dirigentes universitarios, instituciones de investigación y bibliotecarios, plantearon la posibilidad de acceder a los productos de investigación sin cargo para los lectores, buscando nuevas formas de financiar su publicación.

1.2 Surgimiento del acceso abierto

Los diferentes obstáculos que se dieron por parte de las casas editoriales para acceder a la información científica, tuvieron reacciones por parte de los investigadores e interesados en obtener resultados intelectuales útiles en el desarrollo de sus proyectos. Esa respuesta que se promovió con diversas

movilizaciones, declaraciones y propuestas motivó el movimiento por el **acceso abierto** y las iniciativas que a él refieren.

El acceso abierto es un nuevo modelo económico que traslada los gastos de generación de una publicación a las instituciones o a los autores. El movimiento por el acceso abierto es el nombre que se le da a otros modelos de organización, cuyo objetivo es proporcionar el libre acceso y usar el conocimiento científico presentado en forma de artículos, monografías, datos y otros; esto cambia el modelo de publicaciones impresas a otros modelos de financiamiento (Martínez Arellano, F.F., octubre, 2012).

La necesidad de compartir la información provocó que los científicos e involucrados en la ciencia, comenzaran a manifestar sus inconformidades e ideas respecto a la posibilidad de facilitar la comunicación científica, lo que incitó la unión entre ellos, obteniendo así diferentes iniciativas con la finalidad de hacer público el conocimiento científico, resultado de las diferentes investigaciones.

Resultado de lo anteriormente señalado, se puso en funcionamiento (en agosto de 1991) un repositorio denominado Arxiv, el cual contaba con pre-impresos en el área de la física de altas energías y más tarde se amplió a otras áreas relacionadas con las ciencias duras. Con el funcionamiento de este archivo abierto, se hizo evidente que había una demanda en la conservación a largo plazo de la información y de la necesidad de compartir y conocer nuevos conocimientos. Otro hito importante en 1994, fue la propuesta de Stevan Harnad sobre el autoarchivo –sobre el impacto que tendría para la comunidad científica si cada autor depositara sus artículos científicos en un archivo de acceso abierto-. Su existencia fue uno de los factores desencadenantes que llevaron al movimiento actual en la información científica conocida como libre acceso y dio origen a la iniciativa de archivos abiertos.

Como otros ejemplos se tienen las siguientes iniciativas:

Cuadro 1. Iniciativas anteriores al Movimiento por el Acceso Abierto.

AÑO	INICIATIVA	PROPÓSITO
1999	E-Biomed	Hacer el conocimiento y las ideas en ciencias biomédicas amplia y libremente accesibles a la comunidad científica y el público.
1999	Current Science Group	Publicar todos los artículos originales arbitrados y hacerlos disponibles inmediatamente por medio de PubMed Central y su propio sitio web sin cargos por suscripción ni otras restricciones.
2000	PubMed Central	Hacer un depósito de acceso abierto donde se invitaba a los autores a depositar sus trabajos después de publicados.
2000	Public Library of Science (PloS)	Los editores de las revistas científicas tienen legítimo derecho a obtener una recompensa financiera justa por su papel en la comunicación científica. Aunque el registro de las ideas y la investigación científicas no deben pertenecer ni ser controladas por los editores, sino que deben pertenecer al público y deben estar disponibles libremente en línea. Permitir el acceso a los artículos después de un embargo no superior a los seis meses.

Fuente: Sánchez Tarragó, 2009 http://bvs.sld.cu/revistas/aci/vol16_3_07/aci05907.html

Entre las iniciativas que tuvieron éxito y que se publicaron como apoyo al acceso abierto a la información, se encuentran: Declaración de Budapest (Open Access Budapest), Declaración de Bethesda (Bethesda Statement on Open Access) y Declaración de Berlín (Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities).

1.2.1 Iniciativas en favor del movimiento por el acceso abierto

Son varias las iniciativas que se han propuesto defender el acceso abierto a la información, aunque las más aceptadas son tres, -teniendo en cuenta que las recomendaciones son relevantes-. En ellas, se destaca la importancia de establecer políticas que favorezcan la información científica difundida como bien público.

En las tres declaraciones y/o iniciativas, conocidas como BBB o 3B, es definido el acceso abierto y plantean lo siguiente:

1.2.1.1 Declaración de Budapest (Budapest Open Access Initiative)

En diciembre de 2001 el Open Society Institute organizó una reunión en Budapest donde participaron importantes agentes relacionados con los cambios de la comunicación científica, éstos tuvieron influencia en el naciente movimiento por el acceso abierto. La Declaración de Budapest fue firmada por más de trescientas instituciones internacionales; en esta iniciativa se denominó acceso abierto a la forma en que la literatura científica que se encuentra disponible a través de Internet, es gratuita con la finalidad de que cualquier usuario pueda hacer uso de ella como: leer, descargar, copiar, distribuir, imprimir, y con la posibilidad de buscar, enlazar todos los textos de los artículos, recorrerlos para indización exhaustiva, utilizarlos como datos de software o para cualquier otro propósito legal sin barreras financieras, legales o técnicas distintas al fundamental acceso de Internet. La única limitante a la reproducción de los artículos publicados y la única función del derecho de autor, es dar a los autores control sobre la integridad de su trabajo, así como el derecho a ser apropiadamente acreditados y citados (BOAI, 2007).

Asimismo, la definición de acceso abierto de acuerdo con la Declaración de Budapest, contempla que la literatura debe tener acceso libremente en línea, puesto que los académicos la dan al mundo sin la expectativa de recibir pago alguno. Básicamente es la categoría compuesta por sus artículos revisados por pares destinados a publicaciones periódicas, aunque también incluye cualquier *pre-print* sin revisión, que tal vez les gustaría poner en línea para comentar o alertar a otros colegas sobre la importancia de hallazgos de investigación (BOAI, 2001).

La Declaración de Budapest promueve estrategias para hacer frente a los retos de la publicación científica, y encontrar canales más abiertos de comunicación, mediante los trabajos de investigación en depósitos de acceso abierto de

documentos, contemplando así que los documentos principales, para este caso, son los artículos científicos arbitrados.

Del mismo modo, se plantea el impulso a la publicación en revistas especializadas a través de Internet, respetando el derecho intelectual del autor; además esta alternativa promueve aplicar los mecanismos ya establecidos de revisión del trabajo científico para publicaciones de acceso abierto.

Por tanto, para encontrar canales más abiertos de comunicación, se recomienda conseguir que todos los autores publiquen en revistas y hagan uso de archivos o repositorios –de los cuales se hablará posteriormente-, que son una alternativa que tienen para darse a conocer con sus resultados de investigación, y haciendo uso de las rutas o estrategias: la ruta verde y la ruta dorada que sirven como complemento y que se podría conseguir que el acceso abierto fuera total.

Las estrategias o vías verde y dorada complementarias para lograr el acceso abierto, de acuerdo con la Declaración de Budapest, se definen como sigue:

- a) **Ruta o vía dorada.**- Se llama vía dorada al método de publicación de artículos en revistas de acceso abierto; en algunos casos se cobra a los autores por publicar sólo como una forma de sostener la publicación (Hanrad, 2010). Son las publicaciones que se dan a conocer en revistas de acceso abierto (autor paga).

En la ruta dorada las revistas de acceso abierto se diferencian de las tradicionales en sus modelos de negocio en el que ningún coste lo reciba el lector; estas revistas suelen además tener políticas más generosas de reutilización de sus contenidos.

La ruta dorada también ha manifestado un importante avance ya que varios países han empezado a promover la edición en acceso abierto de la literatura científica por medio de revistas de libre acceso.

- b) **Ruta o vía verde** (Archivar en repositorios).- En esta ruta los investigadores realizan el proceso de almacenamiento de sus trabajos de investigación arbitrados, mediante el auto-archivo en un repositorio en este caso, institucional.

También la Declaración de Budapest hace referencia a que las publicaciones que aceptan mantenerse en acceso abierto se pueden ayudar para su financiamiento, con fundaciones y gobiernos que apoyan la investigación, las universidades y laboratorios que emplean investigadores, amigos de la causa del acceso abierto, ganancias de la venta de anuncios incluidos al lado de los textos, fondos liberados de la liquidación o cancelación de revistas que cobran derechos tradicionales de suscripción o acceso, e incluso, con las contribuciones de los propios investigadores. Mismamente, esta iniciativa recomienda no reparar en búsqueda de otras alternativas, pues el objetivo es lograr el acceso abierto para la literatura de revistas doctas arbitradas por pares y obtener una nueva generación de estas revistas.

1.2.1.2 Declaración de Bethesda (Bethesda Statement on Open Access Publishing)

La Declaración de Bethesda fue la segunda reunión con repercusión internacional, celebrada en Bethesda en 2003 y firmada en el Bethesda Statement on Open Access Publishing. Aunque en esta Declaración se defendía el acceso abierto principalmente para las revistas del área biomédica, apoyaba de igual manera a todas las demás ciencias. En ésta se describen dos características con las que deben contar las publicaciones para considerarse como de acceso abierto y éstas son:

- El autor o los beneficiarios del derecho de autor garantizarán a todos los usuarios el derecho de acceso libre, irrevocable, mundial, perpetuo y con licencia para usar, distribuir, transmitir y poner a disposición el trabajo públicamente, hacer y distribuir trabajos derivados, en un medio digital, para un uso responsable, haciendo constar la propiedad del autor, así como derecho a tener copia de parte de las obras para uso personal.

- Que una copia completa del trabajo y de los materiales suplementarios en un formato electrónico localizable normalizado, sea inmediatamente depositada como publicación inicial en al menos un repositorio en línea, que sea mantenido en una institución académica, asociación escolar, agencia gubernamental u otra organización establecida que permita el acceso libre, la libre distribución, la libre interoperabilidad y el archivo a largo plazo (BSOAP, 2003).

El objetivo de la Declaración de Bethesda fue estimular la discusión entre los científicos a nivel internacional relacionados con la biomedicina, sobre la manera de conseguir el objetivo del acceso abierto a la literatura científica.

Entre una serie de consideraciones por parte de la Declaración de Bethesda, se tienen:

1. A los editores les recomienda un compromiso para proporcionar la opción de acceso abierto entre los autores.

2. Cooperar entre editores para desarrollar instrumentos que faciliten la publicación, almacenamiento y búsqueda de manuscritos en formato electrónico.

3. Desarrollar estrategias que faciliten el apego de la iniciativa del acceso abierto para instituciones en desventaja o países en desarrollo.

1.2.1.3 Declaración de Berlín (Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities)

La Declaración de Berlín es otra aportación relevante para promover el acceso abierto, ésta surgió de una reunión en Berlín convocada por la Sociedad Max Planck Institute en el que participaron representantes de varias instituciones europeas, y firmada por 138 instituciones en 2003 que financian proyectos de investigación y menciona:

- ✓ El(los) autor(es) y depositario(s) de la propiedad intelectual de tales contribuciones deben garantizar a todos los usuarios por igual, el derecho gratuito, irrevocable y mundial de acceder a un trabajo erudito, lo mismo que licencia para copiarlo, usarlo, distribuirlo, transmitirlo y exhibirlo públicamente, y para hacer y distribuir trabajos derivados en cualquier medio digital para cualquier propósito responsable; todo sujeto al reconocimiento apropiado de autoría -los estándares de la comunidad continuarán proveyendo los mecanismos para hacer cumplir el reconocimiento apropiado y uso responsable de las obras publicadas, así como ahora se hace-, lo mismo que el derecho de efectuar copias impresas en pequeño número para uso personal.
- ✓ Una versión completa del trabajo y todos sus materiales complementarios, que incluya una copia del permiso del que se habla en el punto anterior, en un conveniente formato electrónico estándar se deposita y así es publicado en por lo menos un repositorio online que utilice estándares técnicos aceptables -tales como las definiciones del acceso abierto-, que sea apoyado y mantenido por una institución académica, sociedad erudita, agencia gubernamental, o bien organización establecida que busque la

implementación del acceso abierto, distribución irrestricta, interoperabilidad y capacidad archivística a largo plazo (DB, 2003).

La Declaración de Berlín recoge las dos condiciones que deben cumplir los documentos de acceso abierto y que se han ido definiendo con más precisión en reuniones continuas.

Igualmente en la Declaración de Berlín se manifiestan las posibilidades que brinda Internet en la difusión del conocimiento y apoya el esquema del acceso abierto, tratando de estimular a los interesados para que los resultados electrónicos de los involucrados en la ciencia sean incluidos en un repositorio que debe ser apoyado por una institución académica, sociedad científica o una organización establecida que se encargue de implementar el acceso abierto, proporcionando distribución irrestricta, interoperabilidad y capacidad archivística a largo plazo.

Con la iniciativas planteadas para lograr el acceso abierto a la información académica, se pretende eliminar las barreras a la información que los investigadores proporcionan sin expectativas de lucro para todos los usuarios de Internet que cuenten con conexión a él, incluyendo los artículos impresos y referidos como pre-impresos (pre-prints), antes siendo revisados por pares que se desee poner a disposición en algún repositorio o base de datos, con la intención de alertar a los colegas sobre descubrimientos importantes pero además, el concepto de acceso abierto implica otorgar permiso para todos los usos académicos de los resultado científicos, es decir, que no solamente elimina las barreras de costo, sino que de igual manera promueve la eliminación de las barreras jurídicas para el uso académico de la información generada por los investigadores, profesores y académicos.

Con las manifestaciones generadas para conseguir el acceso a la comunicación académica se eliminaron barreras, dejando la información libre sin expectativas de lucro y ayudando a los especialistas, en todas las áreas del conocimiento, a estar

alertados sobre nuevos resultados de importancia, de igual forma y como argumenta Martínez Arellano (2011, p. ix), se eliminaron las barreras legales y tecnológicas; para ello, como respuesta se espera mayor accesibilidad a la documentación científica y mayor visibilidad para los autores.

Las BBB promueven dos estrategias para enfrentar los retos de la publicación científica y establecer canales más abiertos de comunicación. Por un lado, se tiene el auto-archivo de los documentos de los investigadores en depósitos abiertos (repositorios) de documentos y por otro, el impulso a la publicación en revistas especializadas de código abierto.

Las tres iniciativas principales (Budapest, Bethesda y Berlín) surgieron con el objetivo común de mejorar el sistema tradicional de comunicación científica y facilitar el acceso a las publicaciones académicas, esta mejora puede concentrarse en tres puntos:

1. Aumentar la rapidez de distribución de los trabajos por medios electrónicos.
2. Reducir los costos de las publicaciones científicas.
3. Aumentar la visibilidad de los trabajos publicados (Melero, Pérez-Agüera, 2004).

Así mismo, los puntos mencionados pueden lograrse a través de diferentes proyectos que pueden resumirse en tres categorías:

- ❖ Creación de repositorios de documentos.
- ❖ Creación de revistas electrónicas de libre acceso.

- ❖ Fomentar y defender el acceso libre a las publicaciones científicas mediante políticas de apoyo a este tipo de proyectos (Melero, Pérez-Agüera, 2004).

Rodríguez Gallardo (2008, p. 67) expone que el objetivo general de las tres declaraciones en favor del acceso abierto, así como proyectos para lograrlo en favor de los resultados de la investigación científica, es optimizar el valor y la calidad de la investigación, y que permita a los investigadores el acceso a la información académica y/o científica bajo mínimas exigencias a través de publicaciones de libre acceso. De esta manera, aunque los objetivos dependan del área a investigar, entre cada área existen coincidencias por la urgencia de obtener la información producida por parte de los diferentes académicos e investigadores en cada área.

1.3 Concepto

El acceso abierto ha sido definido por muchos estudiosos en este tema. Como ejemplo, Peter Suber (2006), define al acceso abierto como aquella literatura digital, en línea, de forma gratuita y libre de la mayoría de los derechos de autor y restricciones de licencia; también hace énfasis en que los contenidos de acceso abierto no se limitan únicamente a los artículos de investigaciones por parte de expertos, porque cualquier resultado de investigación puede ser divulgado siempre y cuando aporte apoyo científico, además ese contenido puede presentarse en cualquier formato de textos o datos, software, audio, vídeo y multimedia.

También menciona Suber (2006, p. 16) que la literatura de acceso abierto presenta las siguientes características:

- Libre disponibilidad de la publicación académica.

- Libre de derechos patrimoniales. El autor otorga licencias mediante Creative Commons.
- Los materiales están disponibles en línea o Internet.
- Los materiales son en texto completo.
- Los materiales son accesibles para las personas desde cualquier lugar sin de discriminación.
- Los materiales pueden ser utilizados libremente por cualquier persona, respetando los derechos de autor.

Por su parte, Martínez Arellano (2011, p. ix) manifiesta que el acceso abierto significa el hecho donde cualquier individuo interesado en esta literatura pueda leer, descargar, copiar, distribuir, imprimir, buscar o enlazar textos completos de los artículos científicos y usarlos con cualquier propósito legítimo, de manera libre y abierta. La única restricción del acceso abierto está en otorga a los autores el respeto a la integridad de su trabajo y el derecho a ser adecuadamente reconocidos y citados.

Ernest Abadal (2012, p. 5) considera acceso abierto a la ciencia, como la disponibilidad libre de los contenidos científicos en Internet, fundamentalmente artículos de revistas. Es decir, que los usuarios no solamente pueden consultarlos gratis sino que además, pueden descargarlos, copiarlos, imprimirlos, distribuirlos, etc., dice que se trata de un modelo de difusión del contenido científico que supone una transformación radical en el funcionamiento de la comunicación científica.

Para el grupo de Redalyc (Red de revistas científicas de América Latina y el Caribe) el acceso abierto es considerado como el nombre de un movimiento

internacional, cuyo objetivo es que cualquier persona en el mundo con una conexión a Internet pueda acceder libremente sin restricción alguna de tipo económico, técnico o legal a la información científica, académica y cultural (Redalyc, 1998).

Por otra parte, muchos autores involucrados en el acceso abierto (como Chalab, Chan, Agengn y Willinsky, entre otros) coinciden en mencionar que el acceso abierto es la forma en la que los interesados obtienen la literatura libre y permanente en Internet con la libertad de usar, copiar, distribuir y adaptar el contenido de acuerdo con las necesidades propias del usuario, siempre y cuando se dé a los autores el control sobre la integridad de sus trabajos y el derecho a ser citados y reconocidos.

Considerando las definiciones anteriores, resulta oportuno concluir que lo que comenzó como una respuesta a la crisis de las revistas científicas, se transformó en un movimiento que exigía el acceso gratuito y sin barreras al conocimiento científico.

Así también, el acceso abierto es definido originalmente por las tres declaraciones y/o iniciativas antes mencionadas, las cuales cuentan con similitudes entre cada una de ellas; todas se derivan de la Declaración de Bethesda en la que se llegó a un acuerdo con los interesados por tener acceso al conocimiento científico de manera libre. Más aún, el movimiento por el acceso abierto surge con fuerza para la edición de las publicaciones científicas, consideradas como uno de los canales más importantes para la comunicación entre los científicos. A partir de la Declaración de Budapest se derivaron dos estrategias para la implementación y desarrollo de ese movimiento: publicación de revistas en acceso abierto o con el modelo antiguo y por otro lado, la creación de los denominados repositorios institucionales en acceso abierto (de los cuales se hablará posteriormente).

El acceso abierto se realiza a través de Internet permitiendo realizar búsquedas, recuperación de documentos, lectura, impresión, distribución y enlace de accesos directos de textos completos de las fuentes, mediante el acceso libre y gratuito.

Melero (2005, p. 256) hace alusión al acceso abierto diciendo que está vinculado con las iniciativas o proyectos que favorezcan y promuevan el acceso abierto, libre y sin restricciones a los trabajos publicados por la comunidad científica; pero es necesario comprender el significado del acceso abierto, ya que es fundamental distinguir entre lo que significa libre acceso (Free Access) y acceso abierto (Open Access), puesto que el primer término significa que es **gratuito** para el usuario final, ésto es que el objeto digital se encuentra en la web y puede descargarse sin pagar por ello; sin embargo, el acceso abierto tiene que ver con los derechos de autor y la posible reutilización de un trabajo **sin ánimo de lucro**; en este caso, el autor determina sus propias licencias a través de alguna de las opciones como Creative Commons –de la cual se hablará posteriormente- además de conseguir el pago correspondiente a las publicaciones así como la infraestructura a utilizar y asignar los derechos y permisos sobre sus trabajos.

Acceso abierto significa la disponibilidad libre de los documentos producidos por todos los investigadores a nivel mundial en cada área del conocimiento, estos textos que están disponibles en Internet pueden ser descargados, distribuirlos, imprimirlos, etc., o simplemente ser leídos, obteniendo por supuesto, la aprobación del autor, quien es el que posee los derechos de cada trabajo para que su distribución sea libre sin depender de editores comerciales que como se ha mencionado, son quienes han puesto las barreras en materia de derechos de autor patrimoniales.

Aunado a lo antes expuesto, el acceso abierto es un contraste agudo del sistema de suscripciones establecido –en el que sólo tienen acceso las personas que son capaces de pagar una suscripción anual-. De este modo, el acceso abierto elimina

las barreras de lo que muchos creen que debería de estar a disposición del público y además prevé la utilización y reutilización de la producción editorial.

Por lo ya mencionado, el acceso abierto tiene las siguientes características principales (Abadal, 2012):

- Es compatible con el derecho de autor intelectual, la revisión por pares, los ingresos, la preservación, el prestigio y otras características y servicios de apoyo asistencial con la literatura académica y convencional.
- No infringe los derechos de autor.
- El acceso abierto se enfoca a los trabajos de los autores que no esperan pago alguno. Estos trabajos son conocidos con el nombre de *Literatura libre de regalías* (Agengn, 2009, p. 12).
- El acceso abierto hace uso de la Licencia Creative Commons.

Aunado a lo anterior, es conveniente indicar que existen lineamientos que establecen las condiciones en las que puede realizarse la reutilización de los trabajos, como ejemplo se tiene Creative Commons, que ofrece permiso de uso a terceras personas en determinadas condiciones:



Reconocimiento.- En cualquier explotación de la obra autorizada por la licencia hará falta reconocer la autoría.



No comercial.- La explotación de la obra queda limitada a usos no comerciales.

 **Sin obras derivadas.**- La autorización para explotar la obra no incluye la transformación para crear una obra derivada.

 **Compartir igual.**- La explotación autorizada incluye la creación de obras derivadas siempre que mantengan la misma licencia para ser divulgadas.

La licencia Creative Commons ofrece al autor, una forma simple y estandarizada de otorgar permiso al público para compartir y usar su trabajo, mediante términos y condiciones. En este sentido, los individuos o instituciones interesados en compartir sus conocimientos, lo pueden hacer con la seguridad de que nadie podrá asumir la titularidad o apropiarse de del intelecto. Con este medio se fomenta la socialización del conocimiento entre un número cada vez mayor de usuarios.

Sumado a todo lo anterior, Martínez Arellano, (octubre, 2012), comenta que el acceso abierto propone libre acceso y resguardo al conocimiento científico en el que entran artículos y publicaciones impresas, que caen de los modelos de pago por suscripción a otros modelos de financiamiento.

El acceso abierto no quiere decir que sea un documento sin costo, sino quiere decir que no se cobra a los lectores o a las instituciones por tener acercamiento a los documentos. Este modelo se creó para hacer accesibles los precios de las publicaciones y no para eliminar sus costos, ya que la creencia de que la información es gratuita es un mito, puesto que producirla siempre tiene un costo (Rodríguez Gallardo, 2008, p. 69).

A esto, Suber (2006, p. 20) considera que el acceso abierto no es gratuito señalando lo siguiente:

Producir la literatura de acceso abierto no es gratis. Ningún defensor del acceso abierto ha dicho jamás que producir esta literatura no tiene costo; si bien, muchos argumentan que es mucho más barata de producir que la literatura publicada de manera convencional e, incluso, que aquella en venta solamente en línea. La cuestión no es si la literatura académica puede hacerse sin costo, sino si existen mejores maneras de pagar el valor monetario.

Rodríguez Gallardo (2008, p. 71) desprende la idea de que mediante la aplicación de algunas opciones de acceso abierto, se ha podido concluir que económicamente es posible su realización, proporcionando al lector un poder para localizar y hacer uso de la literatura relevante, además de que ofrece como ventaja, a los autores y a sus trabajos una vasta y medible visibilidad que les asegura serán leídos y que sus trabajos tendrán impacto.

El movimiento por el acceso abierto promueve el acceso a la literatura científica sin costo para el usuario final, principalmente de los artículos de revistas revisadas por pares (Vinicio Caballero, 2006, p. 95), se eliminan las barreras legales y tecnológicas, tratando de obtener a cambio mayor accesibilidad a los documentos y mayor visibilidad para los autores.

Al promover el acceso abierto entre la comunidad científica, además de cambiar sus sistemas de recompensa, se contribuye a transformar el modelo actual de comunicación científica en un modelo más justo y benéfico para la sociedad.

1.4 Ventajas del acceso abierto

Los resultados de investigación en todas las áreas del conocimiento son un requisito para el progreso, sustentabilidad y bienestar del ser humano. Bajo este supuesto, la literatura científica revisada por pares es un elemento vital para la investigación al estar integrada en el trabajo académico.

Por lo anterior, los científicos siempre han tratado de compartir sus resultados de investigación con los demás. La aparición de Internet aceleró este proceso y posteriormente con las 3B (Budapest, Bethesda y Berlín) al igual que otras propuestas, se hizo posible tener acceso abierto a la información.

Aunque hoy en día existe un debate constante sobre el acceso abierto, como son: las acusaciones de algunos editores tradicionales que implican a la publicación científica de acceso abierto de baja calidad y en peligro (Science News, 2012), se tiene la garantía del sistema de revisión por pares por parte de la comunidad académica, que ha luchado a lo largo de muchos años por conseguir acceso a las publicaciones impresas.

Es posible deducir que existe la incertidumbre de la competitividad o el reconocimiento externo sobre la calidad y el impacto de las publicaciones de acceso abierto.

Existe duda sobre las publicaciones de acceso abierto respecto a que no mantienen lo necesario para convertirse en publicaciones serias, aunque este modelo de publicaciones tienen los mismos estándares de calidad para conseguir el máximo reconocimiento y prestigio.

Al respecto Abadal (2012, p. 5), cita que existen ideas confusas y malas interpretaciones sobre algunos de los objetivos, características o funcionamiento respecto a las publicaciones de acceso abierto; entre estas confusiones menciona las más frecuentes, a saber:

1. Las revistas de acceso abierto, no son de calidad ni sostenibles económicamente.
2. Los repositorios contienen materiales de poca calidad.

3. Los científicos no están interesados ni motivados por el acceso abierto.
4. El acceso abierto tiene presencia residual en la comunicación científica.

Suber (2006, p. 21) defiende al acceso abierto justificando que es compatible con la revisión por pares sumando todas las reglas que tome el acceso abierto para la literatura científica y académica. Al respecto dice también que la revisión por pares no depende del precio o medio de una publicación, el valor, rigor o integridad de la misma.

Continuando con lo anterior, una razón por la que se sabe que la evaluación por parte de pares en las publicaciones de acceso abierto puede ser tan rigurosa y honesta como en las publicaciones comerciales, es el hecho de que pueden usar los mismos procedimientos y normas e incluso de editores y evaluadores que las publicaciones convencionales (Suber, 2006, p. 21); con estas justificaciones hacia las publicaciones de acceso abierto, es posible obtener diversas ventajas, las cuales se pueden agrupar en tres apartados como son:

- Como primer lugar se tiene que el poner los resultados de las investigaciones y de las publicaciones científicas en acceso abierto, tiene como respuesta un mejor funcionamiento en la comunicación por parte de los científicos, ya que se incrementa el uso e impacto de los contenidos, se mejora la calidad de la investigación y se pueden reducir notablemente los costos.
- En segundo lugar, el acceso abierto genera beneficios directos sobre la sociedad, ya que facilita la transferencia directa del conocimiento al entorno económico y social; también es factible disolver las barreras entre los países desarrollados y en vía de desarrollo.

- Por último, es posible decir que con el acceso abierto es factible la reutilización de la información, en tanto que los contenidos ceden algunos derechos de explotación y con ellos se pueden crear productos derivados para investigarlos.

De acuerdo con lo anterior, es posible citar algunas de las ventajas que se tienen mediante el aprovechamiento del acceso abierto.

Las ventajas del modelo de acceso abierto son diversas y se pueden agrupar en tres grandes rubros de acuerdo con Johnson (2012) y Abadal (2012), éstos son:

1. **Ventajas económicas.**- “La puesta a disposición en acceso abierto de los resultados de investigación y de las publicaciones científicas supone una mejora notable del funcionamiento de la comunicación científica, ya que se incrementa el uso e impacto de los contenidos (al estar disponibles sin barreras), se mejora la calidad de la investigación y se pueden reducir notablemente los costos”.
2. **Ventajas sociales.**- Porque el acceso a la información es la clave del desarrollo, la innovación y la prosperidad, como repercusiones sociales importantes se tiene una buena información para los ciudadanos, políticos empresarios y otros. Para todos los países, así como para el público en general el libre acceso a la información, aumentará el conocimiento y la motivación y contribuir con los conocimientos.
3. **Ventaja científica.**- El proceso de investigación se acelera considerablemente debido al acceso inmediato y gratuito a los resultados de la misma, de esta manera la calidad de investigación se mejora pues aumenta la retroalimentación y monitoreo por parte de los investigadores. Así el acceso abierto crea muchas oportunidades, ya que más personas de cualquier área geográfica tratan de resolver los mismos problemas.

Resumiendo, entre las diversas bondades del acceso abierto se pueden mencionar las siguientes:

- ❖ El sistema en línea para la entrega de artículos, la revisión y publicación de los mismos, posibilita un acceso más rápido y directo a los resultados de las investigaciones. Además, se puede hacer pública la versión de un artículo una vez que está aceptado para su publicación sin necesidad de esperar a que sea definitivamente publicado, cosa que puede demorarse algunos meses. Este hecho repercute en la reducción del tiempo existente desde que un documento es publicado hasta que recibe la primera cita.
- ❖ El impacto de la investigación se tiene gracias a la disponibilidad de un artículo en Internet, de esta manera es más probable que tenga más usuarios y citas.
- ❖ Los costos de producción son menos y se proporciona mayor estabilidad que ayuda a los investigadores a tener más control sobre sus obras y conceder una considerable distribución.
- ❖ Universidades y centros de investigación aumentan la demanda de sus resultados a los que llegan sus investigadores y, en consecuencia, sus créditos, ya que sus revistas aumentan su factor de impacto.
- ❖ Los países en su conjunto mejoran su actividad editorial y las citas de la obra de sus científicos y en consecuencia, el crédito de cada país.
- ❖ Se amplía la circulación de las revistas, se da uso amplio por parte de los interesados y por lo tanto se recibe mayor número de citas.
- ❖ Buena promoción para las revistas, es decir, esta forma es la más popular entre los autores, lectores y agencias de financiamiento de investigación.

- ❖ El acceso abierto puede convertir a una revista nacional en internacional.
- ❖ El acceso abierto puede ser una solución para las revistas menos visibles y en consecuencia los autores.
- ❖ Oportunidad para los países en desarrollo ya que pueden acceder a una gama amplia de información desde cualquier lugar del mundo. El acceso abierto rompe las barreras entre países ricos y pobres, pues en los países es variado el nivel educativo, cultural y tecnológico que puede reducirse al mínimo haciendo posible el acceso a los contenidos.
- ❖ Para los estudiantes e interesados es una oportunidad que consiste en revisar literatura de investigación de primera mano.
- ❖ Acceso gratuito a los artículos científicos independientemente de la afiliación a una editorial.

Regularmente las ventajas del acceso abierto van en mayor medida a los países desarrollados, pues los países en desarrollo hacen un esfuerzo para publicar los resultados de su investigación en inglés, puesto que no son hablantes nativos del idioma; por lo tanto, los países desarrollados deben establecer centros de control y análisis para procesar el enorme flujo de información científica con el fin de obtener el máximo beneficio de ella.

1.5 Situación actual

Rodríguez Gallardo (2008, p. 162) refiere que los costos de las suscripciones a las revistas científicas han aumentado en forma sostenida durante más de veinte años, aún por encima de los índices de la inflación mundial y local. Tal situación

ha provocado la búsqueda de formas con el propósito de disminuir los gastos para adquirir la información.

Por su parte Chan (2004, p. 279) plantea que en medida que las publicaciones periódicas se difunden en formato electrónico y se distribuyen en las bases de datos, se encuentra incluido el control de las grandes editoriales comerciales y los usuarios en general, se enfrentan a condiciones de licencias que son cada vez más restrictivas sobre quién puede consultar las bases de datos y además cuándo deben hacerlo, así como el número de usuarios que pueden compartir el material.

Así, existe contradicción referente a los científicos y las organizaciones quienes producen la información y la sostienen para ser publicada –y que en muchos casos proviene de fondos públicos-, dado que no pueden acceder a sus propios trabajos a causa de la resistencia por parte de los editores quienes argumentan que se violan los derechos de autor si se distribuye la información, lo que representa impedimento para obtenerla, siendo que la práctica requiere de artículos para mejorar las investigaciones que llevan a cabo los expertos en cada área y, para ello, existe actualmente un creciente interés y gran preocupación sobre el carácter público de la investigación científica, siendo que numerosas instituciones reclaman que la investigación financiada con fondos públicos, producida bajo el interés público, debe permanecer en el dominio público. Ésto aplica a todos los resultados de investigación, datos y literatura, porque los resultados científicos y técnicos sólo pueden ser posibles compartiendo los resultados de investigación (Sánchez Tarragó, 2009, p. 5).

Con la aparición de Internet se han facilitado los mecanismos de comunicación y, al mismo tiempo, se ha ofrecido la posibilidad de democratizar el acceso de información con servicios alternativos que no cobran por el uso. Por consiguiente, los recursos bibliográficos electrónicos asociados a Internet, han permitido a la comunidad científica hacer públicos los datos de investigación y compartirlos a sus colegas sin tener que cederlos a las editoriales.

La lucha por conseguir el acceso abierto a la información producida por los investigadores y estudiosos ha ganado terreno, por tanto, no se ha dejado de lado seguir modificando políticas para mantener información actualizada vía Internet, así también se han obtenido recursos económicos para seguir manteniendo diversas iniciativas de acceso abierto, como es el caso del The Bipartisan Federal Research Public Access Acta 2006, apoyado por el Congreso de los Estados Unidos donde se destinan recursos para apoyar las iniciativas por el acceso abierto. Por su parte la UNESCO promueve el libre acceso particularmente a la investigación científica, incluyendo los diferentes tipos de publicaciones impresas como: artículos, ponencias, entre otros, que provienen de la investigación financiada con fondos públicos y trabaja junto con sus asociados para dar a conocer los beneficios de libre acceso a los responsables de la elaboración de políticas, investigadores y administración del conocimiento, institutos y centros, a través de su red mundial de oficinas fuera de la sede. Además la UNESCO facilita la elaboración y adaptación de políticas que posibiliten el libre acceso, participa en debates mundiales relacionados con el acceso abierto y coopera con iniciativas locales, regionales y mundiales.

Puesto que la UNESCO promueve el libre acceso a la literatura, incluyendo los diferentes tipos de documentos y publicaciones impresas y trabaja en beneficio del acceso abierto, es necesario recordar las dos vías para la obtención de información de libre acceso incluidas principalmente en Repositorios institucionales o Académicos derivados de la Vía dorada y las publicaciones pertenecientes a un editor comercial o a una Institución sin fines de lucro como las sociedades y aquellas propias de los autores con derechos patrimoniales pertenecientes a la Vía verde.

En la vía o ruta dorada, en la publicación de artículos en revistas de acceso abierto, en algunos casos se cobra a los autores por publicar sólo como una forma

de sostener la publicación. En esta ruta se tienen Revistas de editoriales, instituciones, sociedades, etc.

En la ruta o vía verde, los investigadores realizan el proceso de almacenamiento de sus trabajos de investigación arbitrados mediante el auto-archivo en un repositorio digital. En éstos se incluyen publicaciones propias de un autor con derechos patrimoniales, publicaciones de universidades, sociedades científicas, artículos, objetos de aprendizaje, tesis, literatura gris, apuntes, datos, documentos de texto, sonidos, imágenes, pinturas, entre otros.

Capítulo 2

Repositorios Institucionales

A partir de la Declaración de Budapest son dos estrategias seguidas para la implementación y desarrollo del movimiento por el acceso abierto: La publicación de revistas en acceso abierto o el traspaso del modelo antiguo de publicación, a la edición en abierto y, por otro lado, la creación de los repositorios institucionales o temáticos en abierto. Ambas estrategias son denominadas Ruta o vía dorada y Ruta o vía verde respectivamente. En estas vías, el poseedor de los derechos sobre el trabajo debe permitir la reproducción, uso, distribución y acceso a los documentos a través de Internet con finalidad responsable, es decir, siempre otorgando el reconocimiento del autor.

Una distinción objetiva de un repositorio institucional de acceso abierto, es que se le integran resultados de investigación que se encuentran accesibles y en abierto tales como: tesis, artículos, de revistas, fotos, mapas, documentos económicos o financieros, estadísticas, etc., pero también ideas, propuestas, hipótesis, experimentos, datos, informe de resultados, entre otros. En este sentido, en los repositorios institucionales se aplica la vía o ruta verde, pues los investigadores tienen la oportunidad de integrar sus trabajos mediante el auto-archivo.

2.1 Antecedentes de los repositorios institucionales

Los repositorios institucionales surgieron como respuesta de las instituciones, en especial las académicas, hacia la política inflacionista por parte de las editoriales y la necesidad de conservar, preservar y poner a disposición de su comunidad académica e investigadora su patrimonio intelectual. En este contexto nacieron los

repositorios, que al principio sólo eran archivos digitales pero que fueron evolucionando llegando a formar grandes volúmenes de datos como sonidos e imágenes y que hasta la fecha siguen creciendo en cantidad.

A consecuencia de los diferentes proyectos de automatización que empezaron a tomar fuerza en los años 80 y que se desarrollaron más ampliamente en la década de los 90, surgió la necesidad de promover una accesibilidad mayor a la información bibliográfica, no sólo a los documentos secundarios, sino a los propios documentos primarios.

Un primer intento para el depósito abierto de e-documentos -considerando que un e-documento es la versión digital de un trabajo de investigación impreso- que sirvió como ejemplo para el diseño de repositorios fue el llamado Arxiv –sistema diseñado por Paul Ginsparg en 1991-, con el designio de facilitar la comunicación entre los investigadores (principalmente de especialistas en ciencias físicas); así también se dio acceso abierto para la consulta de pre-impresos, archivos y todos aquellos resultados de investigación que de alguna manera favoreciera a los investigadores. El propósito de este sistema era ofrecer una alternativa que apoyara a la comunicación científica en la que el propio investigador pudiera publicar su producción, o que sus trabajos fueran almacenados en un repositorio, de tal manera que los lectores pudieran conocer sus avances de investigación y recibir comentarios al respecto como retroalimentación.

El éxito del Arxiv sirvió como ejemplo para iniciativas similares en otros campos como el segundo archivo abierto de la Public Library of Science (PLoS) en 2001 que tuvo un seguimiento masivo. Derivado a éstos surgieron otros repositorios más con el afán de hacer uso de Internet intentando proveer una forma alternativa y más barata en el acceso a la literatura, ya sea especializada o general, dependiendo del objetivo del repositorio y, hasta la fecha, se ha visto un aumento masivo en el número de estas herramientas que apoyan la investigación.

Es inevitable considerar que la producción de información siempre ha sido un activopreciado para las universidades por ello, se establecen proyectos de depósitos institucionales gestionados y controlados por las propias instituciones, con la idea de facilitar el acceso abierto a sus resultados y no perder el control sobre ellos.

Los repositorios también conocidos como archivos abiertos o depósitos institucionales no surgieron como alternativa al mundo editorial, sino como un complemento para favorecer la visibilidad de la producción científica y favorecer un valor añadido para la misma, al incrementar el impacto de las diversas investigaciones en las comunidades científicas.

Actualmente se conocen dos tipos de repositorios:

repositorios institucionales.- Son los que recogen la producción de una institución. En este tipo de repositorios, es posible definir políticas para que los interesados accedan a contenidos de la red en la que se encuentran. Son creados por las propias organizaciones para depositar, usar y conservar la producción científica que generan sus investigadores y que ésta se encuentre disponible para el resto de la sociedad, considerando que el conocimiento es un activo más de la organización que debe repercutir en su beneficio.

Repositorios Temáticos o Disciplinarios.- Son aquellos que se crean en torno a una disciplina. Son creados por instituciones en un ámbito del conocimiento para mejorar el sistema de comunicación científica.

Asimismo, es posible hacer mención respecto a que los primeros repositorios fueron temáticos, como ejemplo: Pubmed Central especializado en medicina y biología y en el que se archivan artículos de revistas con la participación de los editores. Más tarde surgieron los repositorios institucionales, que en estos tiempos todavía son promovidos por políticas institucionales que obligan a sus autores al

auto-archivo, por lo que cada centro de investigación está creando su propio repositorio que redundará en el aumento constante de la creación de estos depósitos abiertos.

Frecuentemente se plantea la incertidumbre: si es mejor depositar en un repositorio institucional o en un repositorio temático, cabe decir al respecto que no es excluyente que el documento se deposite en uno, en otro o en ambos repositorios a la vez, puesto que es factible que el autor deposite su documento en una página personal creada al momento de su registro en el repositorio de su institución y en uno temático sin problema alguno.

Los repositorios institucionales se complementan con las revistas de acceso abierto y con la producción investigadora de cada institución u organismo académico.

La creación de un repositorio es la “difusión y aumento del impacto entre la comunidad científica de los resultados de investigación, evitando convertirse en un nuevo depósito infrautilizado y carente de sentido” (Sánchez García, 2007).

En este trabajo es conveniente mencionar que se abordarán los repositorios institucionales de tipo académico, debido a que la preocupación directa es la disseminación de información a investigadores, estudiosos o personas interesadas en algún área del conocimiento, puesto que los repositorios institucionales se comparan con las revistas de acceso abierto, producto de las instituciones a favor de esas investigaciones.

2.2 Características

La producción científica de las instituciones se puede definir como un repositorio que contiene los documentos generados por ellas, éstos son de libre acceso y

tienen permanencia en el tiempo. Ésto, tomando en consideración que la finalidad básica de un repositorio institucional es la difusión y aumento del impacto entre la comunidad científica de los resultados de investigación, evitando convertirse en un depósito no utilizado y carente de sentido.

Los repositorios consisten en tener grandes volúmenes de datos como artículos, tesis, documentos de texto, sonidos, imágenes, pinturas, objetos de aprendizaje, etcétera, en ellos el usuario localiza los documentos y objetos haciendo una descripción de ellos.

Los repositorios se ponen a disposición de los usuarios sirviéndose de una red informática como lo es Internet. Los documentos se almacenan y de la misma forma pueden ser descartados. Tal vez un procedimiento para controlar el ingreso de documentos al repositorio podría ser establecer una política que anuncie el tiempo de permanencia.

Dentro de las características que refieren a los repositorios se pueden anunciar las siguientes: son diversas y muy variadas las capacidades de recuperación de información, no existen estándares y no hay normas específicas que regulen cómo recuperar la información.

Los repositorios proporcionan una serie de servicios a los autores tales como son los datos estadísticos para poder conocer el número de consultas y descargas o también datos para saber desde qué países han sido consultados los documentos, entre otros. Un buen argumento, es la alta visibilidad de los esfuerzos del trabajo científico.

Refiriéndose al contenido de los repositorios institucionales es necesario resaltar que es importante el que se integren documentos institucionales y patrimoniales de la entidad; éstos deben componerse principalmente de los resultados de investigación estrictamente accesibles en abierto. Como ejemplo de esos

documentos que forman parte del repositorio se tienen: tesis, artículos de revistas, fotos, mapas, entre otros; pero también ideas, propuestas, hipótesis, experimentos, datos, etc., siempre y cuando sean el resultado de una investigación correspondiente a la institución en la que fue creada.

El objetivo de la implementación de un repositorio en una Institución es poner a disposición de la comunidad científica y de los usuarios de la red, en forma libre y gratuita, los documentos pertenecientes a la producción científica generada por sus intelectuales para apoyar y sustentar la investigación, mediante la publicación en revistas de acceso abierto y el depósito en repositorios institucionales a los que tengan acceso. De esta manera, es necesario que el contenido del repositorio institucional sea de fácil recuperación y esté disponible en beneficio de la organización y el resto de la sociedad.

Un repositorio o archivo abierto no es un lugar en el que se archivan documentos, sino que obedece a una política preestablecida que indica cómo hacerlo y en qué condiciones.

Los repositorios suponen una ventaja competitiva para la propia organización que proporciona un valor añadido, como lo es el acceso libre a su producción científica a la vez que incrementa la visibilidad y aumenta el impacto propio de sus investigaciones.

Un repositorio institucional cuenta con una política definida que presenta las siguientes características:

- **Auto-archivo.-** Se denomina auto-archivo al procedimiento de depositar los documentos en un repositorio por parte de sus autores para facilitar el libre acceso y gratuito en Internet. Se trata de la segunda vía para conseguir el acceso abierto o ruta verde.

- **Interoperabilidad.**- Uso de procesos normalizados que posibilitan la interconexión con otros archivos abiertos.
- **Libre accesibilidad.**
- **Preservación a largo plazo.**

Los repositorios o archivos abiertos crean y desarrollan archivos digitales del intelecto científico y académico desarrollado para los usuarios, con la característica de ser interoperables y el compromiso de conservarlos mientras la tecnología lo permita, además de acuerdo a las políticas ya establecidas en el momento de la creación del repositorio.

Con la aplicación de Internet y de la World Wide Web se ha visto un aumento en el desarrollo de software libre, lo cual ha dado pauta al surgimiento de los repositorios institucionales. Es por ello que tal vez la implementación de los repositorios en las instituciones tiene un costo relativamente bajo, pues un programa de código abierto puede ser utilizado para tal efecto. Sin embargo, en este aspecto, es necesaria una inversión importante en recursos humanos, sobre todo en la etapa inicial de implementación, además de un servidor potente de acceso rápido. Los costos de desarrollo, pero sobre todo de mantenimiento y conservación, aún no se conocen del todo.

Las tipologías de un repositorio institucional deben basarse en la funcionalidad ésto es, se deben tomar en cuenta los objetivos estratégicos marcados por la institución, la relación que existe entre el acceso abierto y los repositorios institucionales y/o las actividades de organización.

Un repositorio institucional no es solamente un conjunto de software y hardware, sino una serie de compromisos, decisiones y actividades que se necesitan para

administrar los diferentes materiales digitales que se ingresan al mismo, y pensar realmente en preservar la información a largo plazo.

El organismo responsable del repositorio debe encargarse de administrarlo en cuanto a los contenidos y de la distribución de información; así mismo, deberá estar al tanto de los cambios tecnológicos y la migración de contenidos digitales. Igualmente, el desarrollo de repositorios se obliga a contar con la cooperación de diferentes unidades de información interesadas en ello, así como obtener la certificación de la calidad del contenido y notificar la disponibilidad de los contenidos –ya sean publicados, estén por publicarse o, simplemente, formen parte de la literatura gris- a los usuarios finales que cumplan con los requisitos requeridos como son: el proceso de registro que, en caso de los investigadores, podrían obtener un espacio propio que los acreditaría como usuarios que pueden realizar el proceso de auto-archivo. Un requisito indispensable es el que los investigadores conozcan sus derechos y obligaciones (derecho de autor).

Menciona Sánchez García (2007) que “lo cierto y la experiencia lo demuestra, es que no es conveniente incorporar todo tipo de contenido, de e-documentos – considerando que un e-documento es la versión digital de un documento de investigación-, al repositorio. No debe recogerse toda la producción de trabajos científicos electrónicos de una institución, ya que esto conlleva inevitablemente, a la creación de un sistema de almacenamiento”, aunque para evitar este problema, se deben seguir los puntos recomendables para la creación de repositorios que posteriormente se indicarán.

2.3 Definiciones

A continuación se enumeran algunas definiciones similares referentes lo que es un repositorio institucional como sigue:

- Lynch (2003) menciona: repositorio institucional es un conjunto de servicios que ofrece una universidad a los miembros de su comunidad para la organización y distribución de materiales digitales, creados por la institución y los miembros de esa comunidad. Es esencial un compromiso de organización para la administración de los materiales digitales, incluyendo la preservación, así como la organización, acceso y su distribución.
- Por su parte Barton (2005) señala que es una base de datos compuesta de un grupo de servicios destinados a capturar, almacenar, ordenar, preservar y distribuir la documentación académica de una universidad en formato digital.
- Crow (2003) concluye que los repositorios son archivos digitales de la producción intelectual, creados por un grupo de profesores o grupo de investigadores que se encuentra disponible para los usuarios internos y externos.
- De acuerdo con Bustos y González (2007), un repositorio institucional es un archivo electrónico de toda la producción científica producida en una institución, almacenada en un formato digital, que permite la búsqueda y la recuperación para su uso posterior tanto en el ámbito nacional como en el internacional.

Como estos autores hay otros que definen un repositorio de la misma manera pero, concluyendo se puede decir que un repositorio es un entorno en el que se encuentra información digital que ayuda al avance de todas las ramas del saber y, en consecuencia, a todos los estudiosos de una institución académica basado en sus características más importantes y significativas, es el acceso abierto al contenido de los e-documentos y la naturaleza de los mismos como resultado de las investigaciones realizadas por la institución financiadora (Sánchez García, 2007).

De acuerdo con Barton (2005), en los últimos años se ha presentado un movimiento importante, mismo que ha llevado a que los estudiosos consideren la posibilidad de promover repositorios, pues es recomendable que “todas las instituciones de educación superior, usen repositorios institucionales donde se pueda almacenar lo que se publique y donde se pueda consultar de forma gratuita en línea, puesto que la comunicación entre científicos se optimiza reduciendo el tiempo para obtener información actualizada, debido a que los repositorios institucionales representan un avance tecnológico en beneficio de la diseminación de comunicación científica y de libre acceso para todos los interesados.

Los repositorios institucionales son productos definidos por organismos académicos que a pesar de seguir determinadas políticas, son fácilmente adaptables a la estructura de la propia organización.

Los repositorios se desarrollan de acuerdo a las reglas que la propia institución marca, buscando la identificación de los autores y adaptándose para ellos; en los repositorios deberán incorporarse nuevos contenidos producto de investigaciones realizadas por miembros del organismo en el que trabajan.

Los repositorios institucionales entrañan la característica de conservación de la información en cualquiera de sus formatos u objetos a largo plazo, por ello, es posible definir las principales funciones que pueden desarrollarse mediante el ejercicio de:

- La preservación del patrimonio intelectual de la organización.
- El desarrollo de herramientas de búsqueda y recuperación de los resultados de investigación.

- Adaptación de normas para las garantías de los derechos de autor de los investigadores de la institución creadora del mismo.

2.4 Creación de repositorios institucionales

Un repositorio institucional es una herramienta que hace transparente las labores de gestión de una entidad, al mismo tiempo que le da la importancia a la documentación y optimiza los procesos. La conformación de un repositorio institucional permite finalmente garantizar la preservación de la memoria institucional, por ello es de importancia definir qué conservar, cómo hacerlo, qué procesos e insumos tener en cuenta. Estas observaciones son fundamentales para que la tarea de implementar un repositorio, responda a las necesidades institucionales.

En la información de gestión administrativa, es importante poder encontrar el software adecuado que facilite la administración de dicha información, identificando el lugar en la cadena de gestión en que el documento se encuentra y así asegurar que se puede ubicar fácilmente.

El propósito de un repositorio institucional es el desarrollar un entorno rico en información y centrado en la satisfacción de las necesidades de información de los usuarios y, por supuesto organizar, proteger y conservar la memoria institucional. Debe funcionar como un sistema de gestión de contenidos ya que, además de los documentos propiamente dichos, el repositorio ofrece a la comunidad un conjunto de servicios.

El objetivo de la implementación de un repositorio en una institución es: poner a disposición de la sociedad y del resto de investigadores toda la producción científica generada por la institución y hacer que ese contenido sea fácilmente recuperable y disponible en beneficio de la organización y del resto de la sociedad.

Un repositorio institucional no solamente va a incluir artículos de revistas, también es importante que integre documentos institucionales y patrimoniales de la entidad, así como un importante elemento para la recuperación en el aprendizaje puede incluir distintos usos.

Es conveniente recordar que el repositorio no es una alternativa a las revistas, simplemente, es una herramienta que contribuye a favorecer la visibilidad de la entidad y de sus investigadores.

El coste de diseño y puesta en marcha del repositorio institucional depende en muy buena medida de las infraestructuras y recursos humanos disponibles en la propia institución que lo implemente.

Para la creación de un Repositorio Institucional, Barton (2005) menciona seguir lo siguiente:

- ❖ Aprendizaje sobre el proceso, leyendo y examinando otros repositorios institucionales.
- ❖ Desarrollar una definición y un plan de servicio como sigue:
 - Realizar una evaluación de las necesidades de la Universidad.
 - Desarrollar un modelo de costos basado en este plan.
 - Crear una planificación y un horario.
 - Desarrollar políticas que gestionen la recopilación de contenidos, su distribución y mantenimiento.
- ❖ Formar el equipo.

- ❖ Tecnología, es decir, elegir e instalar el software adecuado.
- ❖ Buscar o escoger a quién o a qué institución se le tomará en cuenta para la obtención de los materiales y ofrecerle los servicios.
- ❖ Puesta en funcionamiento del repositorio.

La función de un repositorio institucional es gestionar, proveer el acceso e incrementar la visibilidad de los resultados de investigación de una institución. Barton (2005) afirma que las características propias de un repositorio institucional de libre acceso son las siguientes:

- ✓ El coste de mantenimiento debe ser bajo, lo suficiente para poder cubrir los gastos de publicación electrónica no comercial.
- ✓ El software debe ser gratuito disponible para todos los usuarios puede crearse un software propio o se puede contratar un proveedor comercial, siempre y cuando no sea un software propietario.
- ✓ Las normas de trabajo dentro del repositorio deben estar claramente establecidas y compartidas por todos los implicados.
- ✓ El tiempo de creación del repositorio y la actualización de los documentos incorporados debe ser reducido.
- ✓ Para la difusión, uso eficiente, eficaz y rentable del repositorio, los beneficios obtenidos por el mismo, deben estar claramente identificados, cuantificados y ser medibles.

Es importante definir cómo se usará el repositorio y qué tipo de servicios pretende ofrecer, ya que podría estar interesada la institución en sólo guardar

investigaciones académicas, o quizá incluir sólo tesis de alumnos, materiales docentes, informes universitarios, etc., de la misma manera, debe precisarse si el material será sólo el revisado por colegas o no, si se aceptarán materiales que aún no se publican, entre otras. Así también se deberá definir quién podrá depositar los documentos en el repositorio.

Para planear un repositorio institucional es recomendable:

- Sólo recoger los resultados de investigación con cierto reconocimiento, impacto o visibilidad, pues la calidad en los e-documentos es mucho más importante para garantizar su fiabilidad.
- Un solo repositorio es eficaz dentro de una institución y el uso de los metadatos adecuados determina las funciones a desarrollar.

Una de las primeras cosas que se deben plantear es:

- Si es conveniente crear un repositorio.
- Qué almacenar.
- Si debe contener todo tipo de documentos y éstos deben estar en acceso abierto.
- Pensar en qué posibilidades ofrecen los metadatos en la gestión del repositorio institucional, pues los metadatos permiten el desarrollo de las diversas funcionalidades.

En tal contexto también es recomendable pensar en lo siguiente:

- ✓ Misión del servicio.
- ✓ Tipo de contenidos que se aceptarán.
- ✓ Quiénes serán los usuarios principales.
- ✓ Quiénes son las principales partes interesadas.
- ✓ Si existen recursos ilimitados, qué servicios se pueden ofrecer.
- ✓ Qué se puede ofrecer.
- ✓ Si se cobrará por los servicios*.
- ✓ Responsabilidades de la biblioteca contra la comunidad de contenidos.
- ✓ Principales prioridades de servicio.
- ✓ Prioridades a corto y a largo plazo.

Es recomendable que el diseño de la interfaz del repositorio sea atractiva, fácil de usar y bien documentada para los usuarios; otro concepto clave a tomar en cuenta, es la perpetuidad de sus contenidos, a través de políticas de preservación de sus trabajos. Así se garantiza su acceso mediante la aplicación de estándares de metadatos y de identificación.

Los usuarios serán al fin de cuentas los últimos evaluadores de la funcionalidad real del repositorio, por ello, debe planearse metodológicamente su existencia,

* Mencionan algunos autores que es conveniente ofrecer un servicio y apoyo gratuitos, ya que puede reservarse la opción de cobrar para ofrecer otros servicios de prepago. De esta manera, se llevará de mejor manera un control de costos.

pues al no existir público usuario que lo consulte, el repositorio podría terminar convirtiéndose en un almacén.

Capítulo 3

Repositorios Institucionales: análisis de la situación internacional

3.1 Introducción

La aparición de las TIC trajo importantes cambios para la sociedad en lo que al acceso a la información y al conocimiento se refiere. Inicialmente los soportes físicos, que auxiliaban en la distribución del conocimiento, cambiaron a bases de datos referenciales que indicaban la localización de la información pero no su contenido; ésto no fue suficiente y se pasó a la digitalización de los materiales que formarían parte de repositorios.

El alto costo de las publicaciones, sobre todo de las revistas científicas, estimuló una lucha mundial que buscaba alternativas con la idea de mejorar el acceso al conocimiento científico y eliminar barreras económicas, además de buscar medios que conservaran los resultados de investigación, cuyo resultado fue la creación de repositorios digitales, los cuales tienen como propósito poner a disposición de los interesados los resultados de investigación, no importando las distancias geográficas en las que se encuentren. Así también se cubrió la necesidad de conservar y preservar el patrimonio intelectual.

Entre las estrategias fundamentales del movimiento por el acceso abierto se encuentran: el depósito o autoarchivo en los repositorios de acceso abierto (ruta o vía verde) y las revistas de acceso abierto cuyos contenidos están disponibles gratuitamente en Internet (ruta o vía dorada). Estas colecciones en los repositorios constituyen generalmente archivos digitales que son productos intelectuales de carácter científico y académico, accesibles a los usuarios con pocas o ninguna

barreras económicas y, en cierto modo, de los derechos de autor, con la característica de ser interoperables para que la información contenida en ellos, pueda ser intercambiable entre las diferentes colecciones existentes y son compatibles con la revisión o arbitraje y con los derechos de autor.

Un repositorio institucional se entiende como un sistema de información que reúne, preserva, divulga y da acceso a la producción intelectual y académica de las comunidades universitarias. Entre sus objetivos se encuentran: retroalimentar la investigación y facilitar el acceso a la información científica y académica, así como cumplir con maximizar el uso y el impacto de la misma, en las comunidades investigadoras.

Un repositorio además de almacenar la información posee otras funcionalidades como son: el suministro, gestión, recuperación, visualización y reutilización de documentos digitales. En este sentido, el acceso abierto de un repositorio añade a las ventajas de gratuidad o acceso ilimitado a la información y la fácil disponibilidad de contenidos.

Los repositorios institucionales son archivos digitales que se encuentran accesibles a través de Internet, que reúnen la producción intelectual de una disciplina o de una institución y constituyen una de las estrategias más viables para garantizar el acceso abierto a la información.

Con la propuesta de la Declaración de Budapest (BOAI, 2002), muchas de las instituciones, fundamentalmente académicas, crearon repositorios institucionales por la necesidad de conservar, preservar y poner a disposición de la comunidad docente e investigadora su patrimonio intelectual. En este sentido, actualmente los repositorios generalmente son implementados por instituciones, comunidades temáticas, centros de investigación u otros grupos. En ellos es factible observar que sus contenidos consisten en grandes volúmenes de documentos académicos en diferentes formatos.

3.2 Materiales y métodos

En la comunidad científica se tiene la necesidad de albergar el producto intelectual en forma de colecciones para dar acceso de los resultados de investigación a los interesados, en este sentido, las instituciones reconocen la necesidad de implementar repositorios en los que el resultado intelectual permanezca a largo plazo.

Los repositorios institucionales facilitan el acceso a los contenidos académicos por lo que se autoriza a los interesados a usar eficientemente estas nuevas formas de diseminación ofrecidas por la red.

Los repositorios permiten la exploración y la adopción de nuevas formas de comunicación científica, haciendo uso de medios digitales, los cuales deben ser validados con garantías de accesibilidad a largo plazo, pero también prestan un conjunto de conocimientos que no son publicados, por ello existe la necesidad de implementar estos instrumentos.

Actualmente existe un alto número de repositorios institucionales que permiten el acceso, difusión y utilización de los materiales académicos en torno a la producción de la literatura científica por parte de los investigadores a nivel internacional. Por todo esto, nace la idea de analizar en la base de datos OpenDOAR (Directorio de Repositorios de Acceso Abierto) (Directory of open Access Repositories (<http://www.opendoar.org/>) los repositorios que existen en los diferentes países.

Mediante el análisis se podrán obtener resultados como saber qué países son líderes en la implementación de repositorios institucionales, el software más utilizado y el tipo de documentos que contienen estas herramientas. Así se

apoyará, en este caso, la idea de que los repositorios institucionales cumplen con su propósito original que es el conservar y aceptar documentos exclusivamente académicos en apoyo a la investigación y también la necesidad que existe de contar con estas herramientas por ser parte de las nuevas tecnologías, ya que se debe estar al tanto de las nuevas tecnologías que favorecen la velocidad con que llega la información a manos de los interesados.

Los repositorios digitales sirven como plataforma de almacenamiento, intercambio y difusión de la información científica entre los miembros de la red en el ámbito nacional e internacional, contribuyendo, a la preservación, visibilidad y accesibilidad de los documentos.

Entre los diversos estudios relacionados con los repositorios institucionales, falta mucho por hacer, ya que se tiene la desventaja de no contar con políticas que los rijan para cumplir con lo establecido en el momento de su creación.

El *OpenDOAR* (Directorio de Repositorios de acceso abierto) (Directory of open Access repositories (<http://www.opendoar.org/>) *OpenDOAR* fue creado y es mantenido por la Universidad de Nottingham-. Este directorio fue seleccionado como universo para la el análisis creciente de la situación internacional que guardan los repositorios institucionales, ya que sólo incluye repositorios académicos de acceso abierto.

El Directorio de repositorios de acceso abierto permite la búsqueda por repositorios integrados en general y proporciona una lista de repositorios por continente y sus países, también se puede buscar por idioma, software, tipología de los repositorios y sus estadísticas por cada rubro.

El criterio general para la aceptación de los repositorios –(que se incrementan diariamente) en el banco de datos del *OpenDOAR*-, es recopilar y proporcionar

información sólo sobre sitios comprometidos con el concepto de acceso abierto a los textos completos.

Para este estudio se llevó a cabo un análisis de los repositorios que se enumeran en la base de datos OpenDOAR, con el fin de extraer los elementos necesarios que justifiquen los resultados de este estudio. El análisis se realizó el mes de febrero de 2015.

Los resultados obtenidos se obtuvieron de la siguiente manera:

En primer lugar, se consideraron los siguientes elementos:

1. Proporción de repositorios por continente, pues en todas las regiones del mundo, existen repositorios en mayor o menor medida.
2. Proporción de repositorios por países, se consideró necesario encontrar los países que en mayor medida son creadores de repositorios institucionales y a la vez, conocer cuáles países son los que no cuentan con tantos repositorios.
3. Software, se buscaron las plataformas de software más utilizadas en la implementación de repositorios (libre o propietario).
4. También se buscó la tipología de documentos en los repositorios.

La búsqueda por estos aspectos permitió la recuperación total de 2, 783 repositorios.

En la página del Directorio de Repositorios de Acceso Abierto, se presenta la página de introducción y en ella, se presentan las ligas para buscar el rubro conveniente. Para este estudio se activó la parte de búsqueda (find), en ella se

logró acceder a la parte de búsqueda por algún país (any country); al desplegar este sitio, se obtienen todos los países con el número de repositorios por cada uno de ellos y el continente al que pertenecen.

En el aspecto de software utilizados, en la ventana de búsqueda por algún software (any software), es posible obtener el número de veces que se ha sido utilizado cada uno de ellos en cada repositorio institucional que presenta OpenDOAR. Así mismo, en la ventana de contenido (any content type) se obtiene cada tipo de documentos que de manera general contienen los repositorios integrados al directorio.

Cabe aclarar que OpenDOAR presenta los números totales de cada rubro.

En cada rubro estudiado en este trabajo (repositorios por países y continente, software más utilizado y tipología de los repositorios) se realizó una suma por cada aspecto a considerar; así por ejemplo: para obtener los resultados totales del número de repositorios por continente, se hizo una suma del número de repositorios que tiene cada país para obtener el resultado total. En el aspecto de los software utilizados y tipología de los repositorios, fue más sencillo pues el directorio muestra los resultados totales.

Para las gráficas se realizó la suma correspondiente y los resultados se transformaron a porcentajes para poder representarlos en cada una de ellas.

Mediante la revisión general de los diferentes repositorios que presenta OpenDOAR, será factible observar que los repositorios tienen el fin de mantener y administrar el conocimiento académico y tomar como modelo el software que apunta a la implementación de los repositorios institucionales y, en consecuencia, con el conocimiento que se obtenga de la implementación y desarrollo de los repositorios institucionales, será posible concientizar a las entidades y personas

involucradas directa o indirectamente con las actividades relacionadas en el desarrollo de estas infraestructuras.

Por último y de acuerdo a lo mencionado anteriormente, estos elementos posibilitarán tener conocimiento general de la situación global que guardan los repositorios institucionales a nivel internacional, el software más utilizado en la implementación de los mismos y el tipo de documentos que albergan. Así los repositorios institucionales registrados en OpenDOAR podrán servir de ejemplo para la creación de otros en cualquier lugar del mundo.

3.3 Resultados

Los repositorios institucionales juegan un papel muy importante en la dinámica de preservación y difusión de la memoria institucional y colectiva, de la misma manera, un conjunto de ellos puede considerarse parte de la memoria colectiva de todo un país o cultura.

Universidades e instituciones de todo el mundo han dado marcha a proyectos de repositorios para la organización y difusión de sus contenidos.

Para tener una idea aproximada de la situación internacional que guardan los repositorios, se muestran las estadísticas y gráficas correspondientes que representan los resultados obtenidos del Directorio de repositorios institucionales de acceso abierto (*OpenDOAR*) existentes actualmente en algunas partes del mundo.

A continuación se presentan los resultados obtenidos y se observa cuál es la distribución internacional de repositorios institucionales por continentes y por países, además es fácil saber el contenido de ellos y el software más utilizado.

Para lograr esos resultados se observó primeramente que del total de 194 países que actualmente existen, sólo cuentan con uno o varios repositorios un total de 117 países (38%), pues no todo el mundo ha implementado esta herramienta primaria para los investigadores.

En el directorio de repositorios de acceso abierto *OpenDOAR* reúne, a fecha 4 de julio de 2014, un total de 2746 repositorios repartidos en los cinco diferentes continentes -que en total suman siete regiones (América, Europa, Asia, África, Oceanía, América Central y El Caribe, América del Sur)- actualmente suman como sigue:

Cuadro 2. Resultados globales de repositorios a nivel mundial.

Continente	No. de repositorios
América	543
Europa	1241
Asia	623
África	103
Oceanía	3
América Central y El Caribe	28
América del Sur	242

Total: **2, 783**

Fuente: OpenDOAR, febrero, 2015.

3.4 América

OpenDOAR contempla a Canadá, Estados Unidos y México en América. Estos países suman 539 repositorios.

Canadá cuenta con 64 repositorios, México con 24 y Estados Unidos con 455.

3.5 Europa

Este continente desarrolla un total de 1, 241 repositorios repartidos en sus 35 países que son productivos científicamente.

Sus países son los siguientes:

Cuadro 3. Número de repositorios en Europa.

País	No. de repositorios
Inglaterra	227
Alemania	169
España	115
Albania	1
Austria	15
Belarus	13
Bélgica	28
Bulgaria	6
Croacia	6
Cyprus	4
Checa Rep.	11
Dinamarca	13
Estonia	5
Finlandia	13
Francia	89
Grecia	28
Hungría	24
Islandia	2
Italia	74
Latvia	3
Lithuania	10
Luxemburgo	1
Macedonia	1
Moldova	1
Netherlands	27
Noruega	50
Polonia	85
Portugal	44
Rumania	1
Rusia	22
Serbia	8
Eslovenia	6
Suecia	43
Suiza	17
Ucrania	56

Total: **1, 241**

Fuente: OpenDOAR, febrero, 2015.

3.6 Asia

Este continente cuenta con 31 países y 623 repositorios, de los cuales la mayor parte pertenece a Japón (145) y a la India (68), siguiéndoles Australasia (64), Taiwan (58), Australia (52), China (39), Turquía (45) e Indonesia (38).

Como puede observarse, los países que componen este continente son:

Cuadro 4. Número de repositorios en Asia.

País	No. de repositorios
Japón	145
India	68
Taiwan	58
Australia	52
China	39
Turquía	39
Indonesia	38
Afganistán	1
Armenia	1
Azerbaijan	2
Bangladesh	9
Georgia	3
Hong Kong	2
Irán	5
Irak	1
Israel	2
Kazakhstan	2
Korea Rep.	22
kyrgyzstan	2
Líbano	1
Malasia	21
Nepal	1
Paquistán	3
Filipinas	6
Quatar	1
Arabia Saudita	7
Singapur	4
Sri Lanka	4
Tailandia	9
Viet Nam	1
Nueva Zelanda	12

Total: **623**

Fuente: OpenDOAR, febrero, 2015.

3.7 África

En el Continente Africano se tienen 103 repositorios de los cuales, 29 pertenecen a Sudáfrica.

Los países que pertenecen al Continente Africano son como sigue:

Cuadro 5. Número de repositorios en África.

País	No. de repositorios
Sudáfrica	29
Kenya	12
Algeria	9
Botswana	1
Cameroon	1
Capa Verde	1
Egipto	7
Etiopía	1
Ghana	3
Lesotho	1
Morocco	2
Mozambique	1
Nambia	2
Nigeria	11
Rwanda	1
Senegal	2
Sudán	4
Tanzania	5
Tunisia	1
Uganda	2
Zambia	1
Zimbawe	6

Total: **103**

Fuente: OpenDOAR, febrero, 2015.

3.8 Oceanía

En Oceanía sólo se observan dos países con muy poca implementación de repositorios como son: Fiji con 2 y Nueva Caledonia sólo con 1.

3.9 América Central y El Caribe

América Central el Caribe mantiene los siguientes repositorios:

Cuadro 6. Número de repositorios en América Central y El Caribe.

País	No. de repositorios
Cuba	6
Costa Rica	6
El Salvador	6
Nicaragua	1
República Dominicana	2
Guadaloupe	1
Honduras	1
Jamaica	3
Puerto Rico	1
Trinidad y Tobago	1

Total: **28**

Fuente: OpenDOAR, febrero, 2015.

3.10 América del Sur

En este continente, los 10 países que han implementado repositorios institucionales son:

Cuadro 7. Número de repositorios en América del Sur.

País	No. de repositorios
Brasil	84
Argentina	34
Colombia	35
Bolivia	2
Chile	20
Ecuador	25
Paraguay	1
Perú	24
Uruguay	2
Venezuela	15

Total: **242**

Fuente: OpenDOAR, febrero, 2015.

3.11 Discusión

Es necesario recordar que el acceso abierto es entendido como el acceso inmediato a trabajos principalmente académicos y científicos sin que se requiera suscripción o pago, y que el movimiento por el acceso abierto ha favorecido la creación de repositorios que apoyan para transformar el proceso de publicación de artículos científicos permitiendo el acceso inmediato a las publicaciones arbitradas.

Los repositorios han cobrado importancia en la sociedad académica y científica, ya que representan una fuente de información digital especializada, organizada y accesible para los lectores de diversas áreas.

Los repositorios digitales no tienen un propósito administrativo sino académico, su desarrollo depende de la promoción y el uso que se les dé y para ello, es necesario romper barreras culturales para que los académicos los acepten y se convenzan de su valor para la investigación, en ese sentido, es necesaria la interoperabilidad como forma de apoyar la creación de diversos servicios que se puedan ofrecer a los investigadores e intercambiar la información entre las redes existentes.

Un factor determinante en los países que no apoya el aumento en la implementación de repositorios, es el nivel de analfabetismo.

Paradójicamente los países que más necesitan repositorios con el ánimo de facilitar el acceso libre a la información, son los más pobres aunque por la escasez de recursos debería haber mayor disponibilidad del conocimiento que se genera, para que éste se multiplique más fácilmente.

En el directorio de repositorios de acceso abierto (OpenDOAR, 2015) los repositorios institucionales en Estados Unidos son clara y ampliamente reconocidos como infraestructuras esenciales para la comunicación científica en el

mundo digital; además constituyen infraestructuras de propósito general en el contexto de una práctica académica cambiante.

En Estados Unidos, el administrador sólo veta los artículos por la relevancia para el alcance del repositorio. Los artículos pueden ser depositados en cualquier momento, pero no se hacen públicamente visibles hasta no expire el período de embargo del editor.

Los software mayormente utilizados para la implementación de repositorios son DSpace, EPrint y Bepress y el mantenimiento de ellos se encuentra a cargo de un equipo integrado por diferentes miembros de las comunidades académicas, cada uno de los cuales desempeña una función específica.

En OpenDOAR se puede observar que son pocos los países que dan mayor importancia a los repositorios institucionales, pues en la lista de los 2, 783 repositorios implementados sólo sobresalen: Estados Unidos, Inglaterra, Alemania, España y Japón.

Estados Unidos se presenta en OpenDOAR como el más importante a nivel internacional, pues cuenta con la cantidad de 455 repositorios, en cambio Inglaterra cuenta con 226, Alemania 169, España 114 y Japón 145.

Sin embargo, las regiones más abatidas del mundo ubicadas en África, América Central y El Caribe y Oceanía no cuentan con una vasta representación.

América es un continente que, de acuerdo con OpenDOAR, está formado por tres países que son: Estados Unidos, Canadá y México. Estados Unidos es el país que cuenta con el mayor número de repositorios implementados a nivel mundial (455), siguiéndole Canadá (64) y por último México (24).

Los repositorios institucionales en Estados Unidos y Canadá son reconocidos ampliamente como infraestructuras esenciales para la comunicación científica en el mundo digital. En cambio aunque México sostiene el menor número de repositorios, actualmente forma parte de los países que disfrutan con un soporte jurídico nacional que regula las iniciativas de acceso abierto y repositorios institucionales, con lo cual la información proveniente de fondos públicos es accesible en la web sin barreras legales.

La región europea es una de las más destacadas en la creación de repositorios, ya que los países que la conforman se desempeñan activamente en todas las iniciativas y proyectos relacionados con el acceso abierto, así como en la implementación y desarrollo del autoarchivo.

Europa desarrolla un total de 1241 repositorios repartidos en sus 35 países que son productivos científicamente. Inglaterra es líder en producción científica, y desarrolla iniciativas gubernamentales con respecto al Movimiento por el acceso abierto, en este sentido, cuenta actualmente con 226 repositorios que aprueban el primer lugar en el desarrollo de repositorios institucionales, a ella le sigue Alemania que cuenta con 169.

Inglaterra desarrolla iniciativas para conservar el Movimiento por el acceso abierto. Entre las iniciativas que actualmente se encuentra desarrollando es aquella que alude al establecimiento de una política que impulse el autoarchivo de trabajos científicos financiados con fondos públicos.

Inglaterra es el tercer país después de Estados Unidos y Alemania con mayor cantidad de repositorios establecidos en universidades.

En Alemania, es posible decir que es uno de los países con producción científica relevante a nivel mundial y, al igual que Inglaterra, es líder en las iniciativas en favor del acceso abierto, pues cabe recordar que en este país se desarrolló la

Declaración de Berlín que marcó pautas para el desarrollo del Movimiento en favor del acceso abierto.

Sin embargo, España ocupa el tercer lugar en la implementación de repositorios en el Continente Europeo, ya que sólo cuenta con 114 repositorios.

Resumiendo, el continente Europeo es uno de los más destacados en la implementación de repositorios institucionales, pues los países inmersos muestran un desempeño activo en todas las iniciativas y proyectos relacionados con el acceso abierto y la implementación y desarrollo de la práctica del autoarchivo.

El autoarchivo se lleva a cabo en las diferentes instituciones por los propios investigadores, es decir, la producción científica institucional se alberga completamente en los repositorios por los mismos investigadores.

En cuanto al Continente Asiático, es posible decir que posee 33 países y 620 repositorios de los cuales, la mayor parte pertenece a Japón (145) y a la India (68), siguiéndoles Australasia (64), Taiwan (58), Australia (52), China (39), Turquía (39) e Indonesia (38).

Japón se encuentra entre los países líderes en la producción científica mundial, pues junto con Estados Unidos, Inglaterra y Alemania, produce un porcentaje considerable en las publicaciones más citadas en el mundo. Además fue uno de los países que junto con Brasil y Finlandia participó en el proyecto de licencias Creative Commons, lo que muestra su interés en legislaciones e iniciativas que apoyan el acceso abierto a la información (Flores Cuesta, Sánchez Tarragó, 2007).

Por su parte la India está a la vanguardia del resto de los países subdesarrollados, tanto en términos de crecimiento económico como de productividad científica, además es líder en la creación de repositorios institucionales.

Refieren Suber y Arunachalam que existen estadísticas y estudios referentes al uso de algunos repositorios Hindúes, lo cual muestra que investigadores de naciones desarrolladas se encuentran accediendo a la literatura disponible en estos archivos y en revistas de acceso abierto por lo que la India se encuentra liderando el movimiento haciendo que los países subdesarrollados descubran la literatura científica de calidad que producen estos países (Suber & Arunachalam, 2006).

En África se cuenta con 102 repositorios de los cuales 29 pertenecen a Sudáfrica. Aquí se puede apreciar que a pesar de ser un continente subdesarrollado, existe un creciente interés por mantener repositorios académicos que mantienen su información actualizada y, que a la vez, están en favor del acceso.

En la región Africana, a excepción de Sudáfrica, el resto de los países tienen de 1 a 12 repositorios. Esto demuestra que la escasez de recursos y su disponibilidad están en iguales condiciones.

En Oceanía se carece de repositorios pues al contar sólo con dos países se tienen sólo 4, tres en Fiji y 1 en Nueva Caledonia.

América Central y El Caribe tiene un total de 28 repositorios, tal vez por tratarse de un continente subdesarrollado no es posible mantener nuevos repositorios aunque no se descarta la idea de futuros nuevos repositorios.

En América Central y El Caribe sobresale Cuba, Costa Rica y El Salvador, en cambio los demás países sólo cuentan con uno o dos repositorios.

En América del Sur existen 239 repositorios, la mayor cantidad de ellos pertenecen a Brasil (84) y le continúan: Argentina y Colombia con 34 cada uno.

Es posible que en Brasil se concentre el mayor número de repositorios debido a que en este país surgió el proyecto Scielo a través del cual se desarrolló una metodología para preparar, almacenar, disseminar y evaluar la literatura científica en formato electrónico.

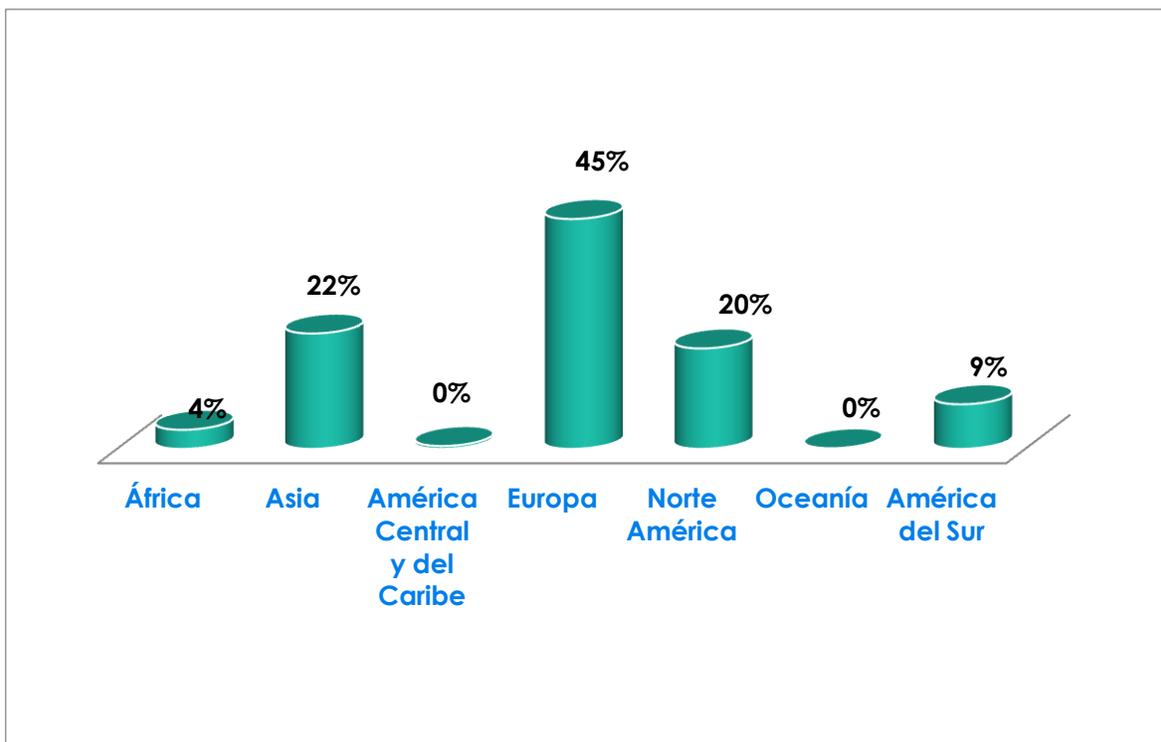
Los materiales o documentos depositados en los obedecen a políticas y prácticas que resguardan los derechos de quienes depositan documentos en el repositorio. Quien deposita información conserva todos los derechos de propiedad intelectual sobre el material. Los únicos derechos que se poseen en el repositorio sobre los documentos que se depositan, son los de hacer copias para respaldos o, el de migrar colecciones a otro formato, en caso de ser necesario.

Por lo anterior, es necesario decir que crear infraestructuras, como son los repositorios institucionales, de interoperabilidad entre las comunidades académicas científicas y culturales que incorporen estándares de cambio, se está convirtiendo en uno de los proyectos trabajados por varios países. De igual manera, es posible mencionar que el objetivo de las redes de alta tecnología se está cumpliendo, ya que existe la conexión entre universidades y centros de investigación para promover el intercambio de información.

3.12 Proporción de repositorios por continente

El aumento diario de repositorios se concentra en los continentes más desarrollados científicamente, tal como se puede observar en la siguiente gráfica según datos obtenidos de OpenDOAR.

Gráfica 1. Proporción de repositorios por continente



Fuente: OpenDOAR, febrero, 2015.

En relación a las regiones de la gráfica 1, se puede apreciar que el Continente Europeo es el más productivo a nivel internacional en la implementación de repositorios institucionales.

En Europa los repositorios suman el 45% del total de repositorios a nivel mundial de su producción, siguiendo a este continente la región asiática con el 22%, en el que gracias a Japón y la India se ha podido llegar a este resultado. Sin embargo, Norte América ocupa el tercer lugar dentro de la gráfica (20%), gracias a que Estados Unidos tiene un total de 451 repositorios.

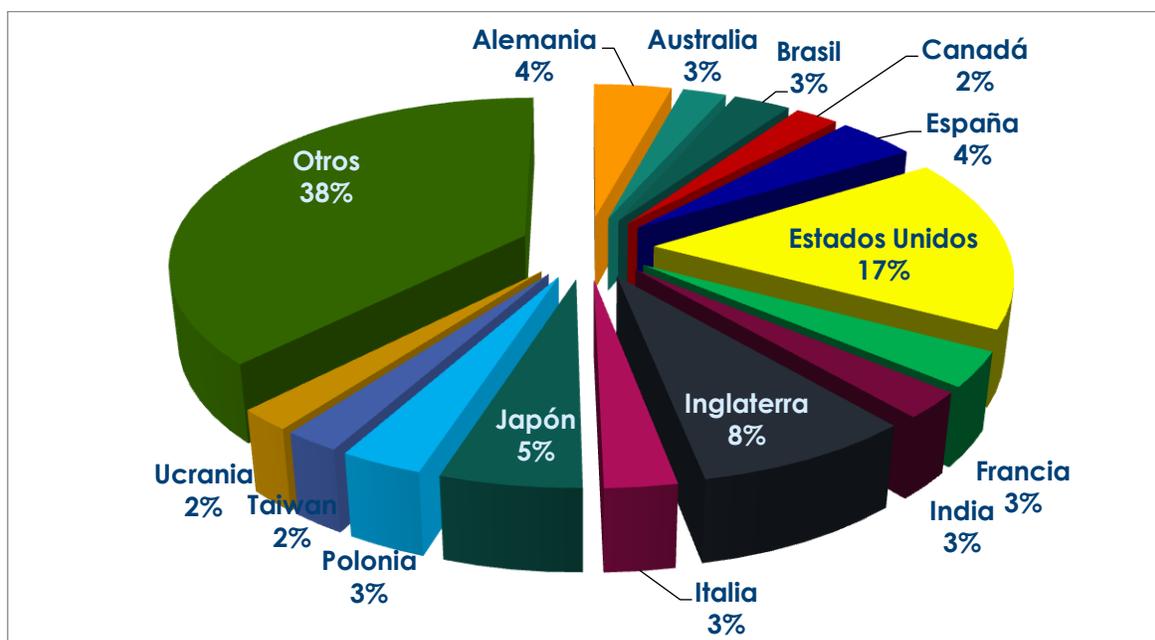
África (4%) alcanzó a representarse, pues la implementación de repositorios en este continente es muy poca; en cambio Oceanía y América Central y el Caribe no tienen representación en este estudio a consecuencia de su poca producción,

pues América Central y del Caribe suman un total de 28 repositorios (entre América Central y El Caribe) y Oceanía sólo tiene 4 repositorios.

3.13 Proporción de repositorios por países

Un aspecto interesante que se puede apreciar en la gráfica 2, es el hecho de que dentro de los países con más repositorios institucionales registrados en OpenDOAR, sólo Estados Unidos se destaca con el 17% del total de repositorios analizados y los demás países sólo contemplan entre el 8% y el 3% del total.

Gráfica 2. Proporción de repositorios por países.



Fuente: OpenDOAR, febrero, 2015.

Sólo 14 países (7%) figuran en las estadísticas representadas en la gráfica, del total de 194 países que se contemplan en OpenDOAR.

2,746 repositorios son los que conforman el 100% y de ellos sólo 1,812 (40%) son los que figuran en la gráfica, repartidos entre 14 países, entre ellos se encuentran

los siguientes países: Inglaterra con el 8%, Japón suma el 5%, Alemania tiene el 4% al igual que España. Australia, Brasil, Francia, India, Italia y Polonia tienen el 3% cada uno; Canadá, Taiwan y Ucrania sólo cuentan con el 2%; sin embargo, los países menos productivos reúnen un total del 38%, pues si se deseara figurarlos en la gráfica no aparecerían.

3.14 Proporción de software utilizado en repositorios

El software juega un papel fundamental en la construcción, estructuración y el mantenimiento de Repositorios Digitales.

Existen diferentes software para la creación de repositorios institucionales de acceso abierto tales como: Dspace, Fedora, Eprints, Greenstone entre otros.

En los resultados obtenidos en OpenDOAR, se encontró que los software utilizados en los repositorios a nivel internacional fueron: CONTENTdm, Digibib, Digital Commons, DigiTool, Diva-Portal, dLibra, Drupal, DSpace, Eprints, ETD-db, Fedora, Greenstone, HAL, HTML, invenio, Open Repository, OPUS y Scielo.

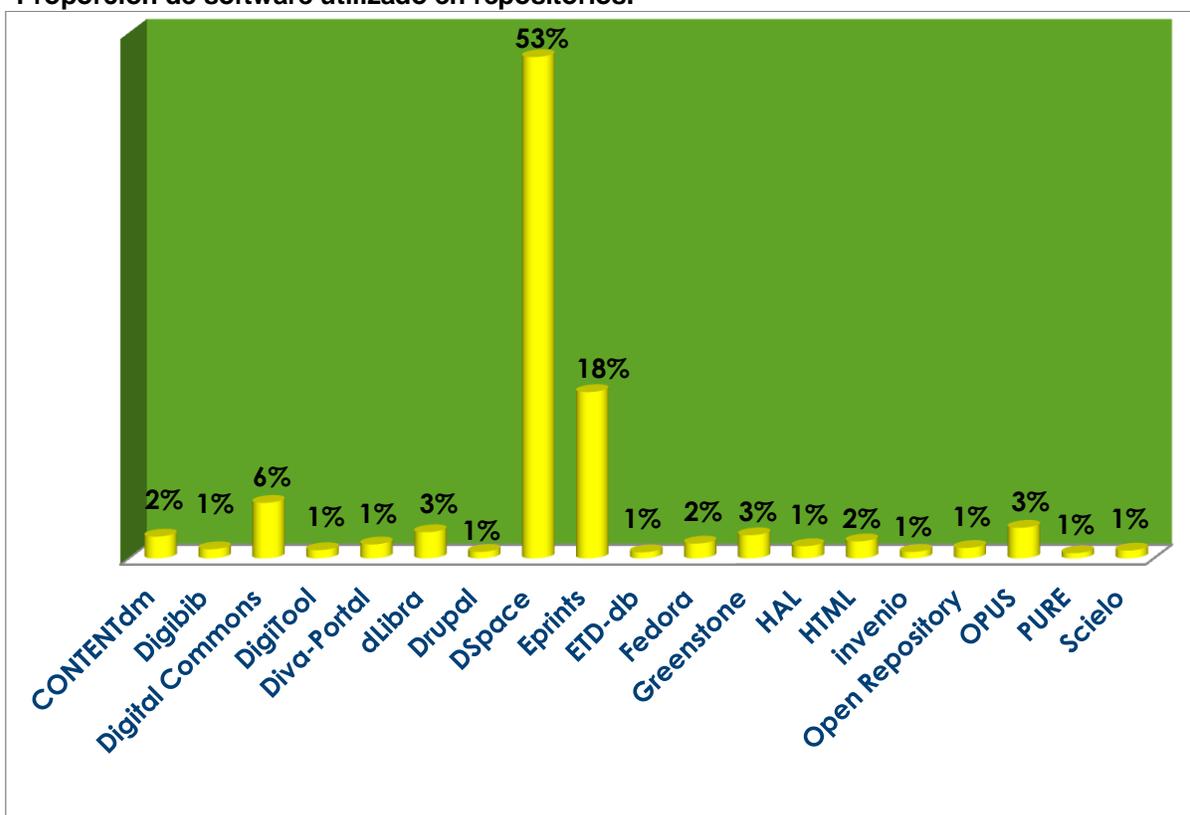
HTML (HyperText Markup Language) es un lenguaje de marcado que sirve como estándar de referencia para la elaboración de páginas Web en sus diferentes versiones, define una estructura básica y un código para la definición de contenido como texto, imágenes, entre otros.

Scielo (Scientific Electronic Library Online o Biblioteca Científica Electrónica en Línea) es un proyecto de biblioteca electrónica -iniciativa de Fundação de Amparo à Pesquisa do Estado de São Paulo, Brasil, la Fundación FAPESP y el Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud- para el apoyo de la investigación que permite la publicación electrónica de ediciones completas de las revistas científicas mediante una plataforma de software que

posibilita el acceso a través de distintos mecanismos, incluyendo listas de títulos y por materia, índices de autores y materias y un motor de búsqueda.

Por lo anterior, cabe aclarar que HTML y Scielo no son software pero debido a que son mencionados como herramienta para hacer funcionar repositorios institucionales, es conveniente tomarlos en cuenta como tales, de cualquier forma, cumplen con la función de poner en marcha algunos repositorios institucionales de acceso abierto contemplados por OpenDOAR.

Proporción de software utilizado en repositorios.



Fuente: OpenDOAR, febrero, 2015.

En la revisión hecha en *OpenDOAR* se obtuvo que 133 veces fue utilizado un software, aunque sólo 19 software fueron realmente utilizados en la implementación de los distintos repositorios, algunos son de código abierto como: Diva-Portal, DLibra, Drupal, DSpace, EPrints, ETD-db, Fedora, Greenstone, HAL,

Invenio, OPUS y Pure; pero otros software son propietarios como: CONTENTdm, Digibib, Digital Commons, DigiTool y Open Repository.

Es posible observar en la gráfica 3, que las plataformas de software más utilizadas son: DSpace con 1145 veces en varios repositorios, lo que representa el 53%, a éste le sigue EPrints con 379 (18%) aplicado a diferentes repositorios y por último Digital Commons con 126 usos (6%) y OPUS con 71 repeticiones (3%).

A pesar de no contar con precios de implementación de repositorios en el *OpenDOAR*, se percibe que el costo de cada repositorio es bajo tal vez por la existencia de una gran cantidad de plataformas de software libre que permiten la gestión de documentos digitales.

3.15 Tipología en los repositorios institucionales

Tipología en los repositorios institucionales.



Fuente: OpenDOAR, 2015.

Destacan también en la tipología de documentos depositados en los repositorios institucionales registrados en OpenDOAR, los artículos generados por una institución o son publicados en revistas, los cuales suman el 22% (1853 artículos), las tesis con el 17% (1425 tesis), los libros 12% (1008), trabajos no publicados 12% (985 trabajos no publicados), objetos de aprendizaje 5% (452 objetos de aprendizaje), las referencias 5% (429 referencias), los trabajos especiales 6% (471 trabajos especiales) y (las conferencias 11% (959).

3.16 Conclusión

La comunicación científica es un mecanismo que ayuda a la difusión de los resultados de investigación entre los investigadores; es imprescindible en el desarrollo de la práctica de la ciencia y está presente en todas las etapas del proceso de investigación.

En la era de las redes y tecnologías de información, favorecer la comunicación más abierta posible a los resultados de investigación, ha traído como resultado la reflexión y compromiso de científicos, instituciones y gestores de información, en el aspecto de la calidad, visibilidad e impacto de los resultados de científicos, la gestión de los derechos de autor y la garantía de preservación digital.

El arribo de las tecnologías de información y comunicación, además de la Web, han modificado la manera en que se divulga la información que pretende ser preservada y han propiciado una evidente mejora en los canales de acceso a la información; el factor fundamental en este proceso ha sido el cambio de soporte del formato tradicional impreso al digital.

Una de las propuestas manifestada como resultado del Movimiento por el acceso abierto a la información científica, se ve concretada en la creación de repositorios institucionales en acceso abierto en todo el mundo, éstos han ido aumentando aceleradamente. En OpenDOAR es posible observar un panorama mundial de estos repositorios que se concentran más en los países desarrollados como Estados Unidos e Inglaterra. En los países con menor desarrollo como Fiji y Puerto Rico, los repositorios se tienen en menor cantidad, esto refleja la falta de cultura académica, esto conlleva a que no existe la preocupación de avanzar tecnológica y académicamente, lo cual resulta evidente que sin la colaboración de los autores no se alcanzarán los objetivos del acceso abierto y tampoco será fácil justificar la creación y mantenimiento de los repositorios institucionales.

La falta de conocimiento y el desinterés por parte de los científicos hacia los repositorios digitales detiene la implementación de éstos y el desarrollo a nivel mundial.

Muchos de los investigadores no le toman importancia a los Repositorios Institucionales o mantienen resistencia al hacer uso de estas infraestructuras, entre los factores relacionados, se pueden mencionar los siguientes: desinterés por este tema, desconocimiento de los derechos de autor, resistencia al cambio, desconocimiento de los objetivos de estos sistemas, falta de tiempo o recursos, indiferencia hacia los posibles beneficios, resistencia a este tipo de publicación abierta, pérdida del control sobre sus trabajos, autor, la desconfianza al plagio, entre otros, es así que mientras los repositorios no cuenten con una revisión por parte de sus pares seguirá existiendo falta de conocimiento sobre ellos y falta de credibilidad. Desafortunadamente es necesario todavía romper diversas barreras culturales para que los académicos acepten, reconozcan y se convenzan del valor que tienen los repositorios para la investigación y mientras sigan existiendo los peros a los repositorios, será difícil el desarrollo un país.

Para dar solución al desconocimiento de los repositorios institucionales por parte de los investigadores es necesario convencerlos de los beneficios que se tienen con esta infraestructura, así como dar entrenamiento para su uso. Como ejemplo se tiene que algunas universidades para favorecer la política del autoarchivo de sus investigadores, recomiendan a sus científicos que se ocupen únicamente de depositar sus documentos y los permisos los gestiona la institución así esta ayuda sirve como soporte inicial a los autores para que depositen sus documentos, de manera que puedan convencerse de la facilidad y poca inversión de tiempo que conlleva; También algunas instituciones han establecido una política de autoarchivo, obligando a depositar una copia de los trabajos científicos en el repositorio financiado por la institución.

En los repositorios institucionales se encuentra una colección de objetos digitales basada en la web y constituida por material producto de resultados de investigación obtenidos por los miembros de una institución. Estos repositorios presentan diversas características entre las cuales se pueden mencionar: el autoarchivo, la interoperabilidad con otros sistemas para que la información contenida pueda ser intercambiable entre las diferentes colecciones existentes y la preservación a largo plazo. Entre sus ventajas se tienen: mayor citación de los trabajos concernientes a los investigadores, mayor rapidez de la publicación, visibilidad, entre otros.

En sus inicios, el objetivo de los repositorios fue acelerar el acceso libre a los resultados de investigación, concretados en artículos científicos, ponencias, presentaciones o libros. Luego se fueron incluyendo otros materiales y hasta objetos. Sin embargo, conceptualmente se diferencian de las bibliotecas o hemerotecas virtuales y de los portales en la web pues los repositorios no consisten en una simple colección de documentos y objetos digitales, sino que adquieren límites más definidos para alcanzar los objetivos propuestos, sobre todo el principio de libre acceso al conocimiento que es generado con fondos públicos.

Entre los repositorios que presenta OpenDOAR se tienen diferentes tipos de documentos como son: artículos, trabajos no publicados, trabajos especiales, tesis, software, referencias, grupos de datos, objetos de aprendizaje, multimedia, libros, patentes y conferencias.

Entre las plataformas de software más populares disponibles actualmente para la creación de colecciones digitales se encuentran DSpace, Greenstone, Fedora y CONTENTdm para la creación de repositorios, pero como se muestra en OpenDOAR, se puede emplear cualquier otro software como: Digibib, Digital Commons, Diva-Portal, dLibra, Drupal, EPrints, ETD-db, HAL, Invenio, Open Repository, OPUS, Pure, entre otros, ésto será dependiendo de las necesidades de cada repositorio. Cabe aclarar que un repositorio debe ser lo más económico posible y es recomendable hacer uso de un software de acceso abierto que permita la introducción de todos los trabajos académicos no importando el tipo o forma.

Capítulo 4

Metadatos y software para repositorios institucionales

La evolución de la web ha convertido las páginas y documentos digitales estáticos por documentos generalmente dinámicos, en este sentido, se ha cambiado del concepto de publicación de páginas web simples a esquemas más complejos. En los repositorios institucionales se maneja información masiva en diversos formatos en todas las disciplinas y este avance los favorece.

La web consiste en una red de páginas con enlaces hiper-textuales que contienen todo tipo de recursos, estas páginas se encuentran en diferentes servidores. Los usuarios usan programas para poder visualizar el contenido de los recursos y seguir los enlaces para acceder a diferente información en múltiples formatos. La web constituye un inmenso repositorio de documentos, imágenes, sonido, y más, puestos al alcance de los interesados, en este caso, los metadatos juegan un papel muy importante para la recuperación de documentos digitales denominados recursos.

La evolución de la web exige que los metadatos alcancen nuevas áreas de aplicación en donde juegan un papel primordial principalmente para localizar recursos, además constituyen un conjunto de elementos cuya semántica representa la información electrónica e implica su descripción de manera conjunta al recurso aunque sea almacenado en otro lugar.

En el contexto dinámico que ofrece la tecnología digital, surgió la iniciativa de archivos abiertos como alternativa para solucionar inconvenientes que presentan los sistemas de información inmersos, esta iniciativa desarrolla y promueve la interoperabilidad entre diferentes medios de distribución de información que tienen

como objeto facilitar la difusión de los contenidos, constituyendo así una herramienta que, gracias al Movimiento del Acceso Abierto, fue impulsado por científicos y especialistas de la información con la finalidad de poner al alcance de toda la sociedad los contenidos científicos y académicos.

En el contexto global y cambiante que ofrece Internet, los espacios de registro y almacenamiento de información académica y científica -como son los repositorios institucionales- presentan el desafío de integrar herramientas a los objetos digitales como son los metadatos, con el objeto de hacerlos visibles y dinámicos.

Los repositorios aumentan su valor cuando evidencian el impacto de sus contenidos, por lo que es fundamental ofrecer metadatos precisos y completos que permitan que los objetos depositados sean identificados, localizados y reutilizados.

Los repositorios presentan conjuntos de metadatos. Los metadatos pueden estar estructurados en forma de conjunto de elementos previamente definidos, es así que las diferentes necesidades de cada proyecto han definido conjuntos de ellos – conocidos como sistemas o esquemas de metadatos-, pensados para aplicaciones concretas y como consecuencia han aumentado en su número. Esta situación conlleva a pensar que cada entorno de trabajo tiene un esquema de metadatos propio y con ello se pretende interactuar con otros grupos, aunque para ello es necesario contar con un nivel semántico en cada uno para producir un intercambio que positivamente en la relevancia de los resultados de búsqueda.

Mundialmente se han creado infinidad de recursos informativos digitales en diversos formatos dirigidos a todos los sectores de la sociedad; sin embargo, en la mayoría de las ocasiones pasan inadvertidos y no son aprovechados debido a las carencias de metadatos, es por ello de gran importancia contar con un software especializado que permita la edición de éstos.

La mayor parte de los sistemas disponibles para la implementación de repositorios institucionales son de software libre el cual es el más recomendado por los expertos. Además de ser gratuito en su adquisición, generalmente depende de otras tecnologías de consideración libre como Java, servidores Apache, bases de datos MySQL, entre otros y se basa en estándares internacionalmente aceptados en cuanto a formatos de archivo y esquemas de metadatos.

El software libre ofrece estabilidad y flexibilidad en su manipulación permitiendo la implementación de editores de metadatos. Se compone de un código fuente que permite interoperar y potenciar el uso de la información independientemente del tipo de de los materiales digitales existentes en las instituciones.

Por ser herramientas fundamentales para la implementación y la puesta en marcha de un repositorio institucional, es indispensable señalar el tipo de metadatos más utilizado y que se acomoda a todo tipo de materiales para su descripción, así también describir algunos tipos de software que por su estructura y gratuidad, favorecen la recuperación de la información.

En este estudio se analiza un esquema de metadatos seleccionado entre la gran cantidad de esquemas ya establecidos. Esta preferencia se debe a que se considera el más completo y fácil de aplicar para la descripción de cualquier tipo de material, además tiene relevancia y popularidad. Se describe además porque la mayoría de los repositorios analizados, enumerados en openDOAR lo utilizan. Se trata de Dublin Core usado principalmente en el desarrollo de repositorios institucionales y bibliotecas digitales, ya que gran parte de software se apegan a este esquema.

Por otro lado, se presentan tres tipos de software que por un lado son los más utilizados y cumplen con las características básicas como: garantizar la interoperabilidad, son de código abierto y están disponibles en forma pública y gratuita para su descarga e implementación. Los tres software que se tratan en

este capítulo son los siguientes: Dublin Core, EPrints, y Greenstone. Los dos primeros, son herramientas que se utilizan para el desarrollo de repositorios, principalmente académicos, en cambio, el tercero es para el desarrollo de bibliotecas digitales.

4.1 Metadatos

El mundo está inmerso en una explosión de información, este cambio se ha visto potenciado por la aparición y especialización de nuevas tecnologías de información que lleva a retos y desafíos imposibles de ignorar.

Internet y especialmente la WWW (World Wide Web), han traído cambios en el paradigma de organización del conocimiento.

Una de las herramientas creadas con el fin de tomar control, organización y descripción para la información digital son los llamados metadatos, con los cuales se ha logrado un avance continuo, logrando brindar soluciones sencillas y muy completas para el caos que existe dentro de la información en Internet.

Refiriendo que la aplicación de los metadatos es para un ambiente digital, se considera que el procesamiento documental de los objetos digitales a través de metadatos, no se puede llamar **catalogación** en el sentido clásico del concepto a causa de que la profundidad y los aspectos que pueden llegar a manejar los metadatos, van más allá de la simple descripción documental y temática propuesta por reglas como las Reglas de Catalogación Angloamericanas o las ISBD; debido a que la información digital exige un tratamiento diferente al de los materiales impresos, porque sus protocolos y estándares son simplemente para transmisión de datos.

La catalogación difiere de los metadatos por tratar los distintos materiales con una descripción general, mientras que la asignación de metadatos implica soluciones distintas y suplementarias a la hora de describir los documentos entendidos como objetos de información en la Red.

Murtha Baca (1999, p. 7) desprende la idea de que los metadatos son importantes en el proceso de recuperación de información digital por lo siguiente:

- ✓ Proporcionan **más accesibilidad**, pues mejoran la efectividad en la búsqueda, además facilitan la recuperación de la información dado que pueden actuar en diferentes niveles, es decir, en la colección u objeto mismo.
- ✓ **Retención del contexto** porque mantienen relacionados los objetos digitales a los documentos reales con mayor exhaustividad.
- ✓ **Uso más amplio**, pues facilitan el acceso a usuarios ubicados geográficamente en diferentes lugares.
- ✓ **Versiones múltiples** al interrelacionar las diferentes versiones de un documento digital y lo relaciona con el objeto primario.
- ✓ **Cuestiones legales**, pues permite a los depositarios identificar los niveles de información sobre los derechos y las restricciones que afectan al documento digital en sus diferentes versiones.
- ✓ **Conservación**, tomando en cuenta que los metadatos garantizan que los documentos digitales perduren independientemente de la evolución de las tecnologías.

- ✓ **Mejora de sistemas y economía**, ya que la información técnica que proporcionan, permite evaluar los sistemas y corregir sus errores.

El crecimiento incontrolable de los recursos de información existentes en la web, han provocado que los usuarios se enfrenten a una red confusa, carente de un orden necesario para recuperar la información pertinente. En este sentido, los profesionales de la información e interesados en ella, han visto modificadas sus tareas buscando la forma de organizar la información en Internet para garantizar su localización y recuperación exitosa.

El procesamiento y manejo de los recursos digitales tienen como pilar a los metadatos, cuyo uso normalizado se presenta como solución a los gestores y creadores de recursos de información.

El término **metadatos** es anterior a la Web y fue acuñado en los años sesenta del siglo XX por Jack E. Myers para describir “conjunto de datos”, aunque la palabra apareció escrita hasta 1973 (Méndez Rodríguez, 2002). En 1980 el término comenzó a aparecer frecuentemente en la literatura sobre sistemas de gestión de bases de datos.

Más allá de la identificación terminológica, etimología y evolución de los metadatos, lo destacable es que el concepto estricto surge en la web.

La tipología, funciones y modelos de metadatos, presentan múltiples alternativas que dependen de factores también diversos como: el tipo de información que describen, el nivel de estructuración, el lugar donde se encuentren, su ámbito de aplicación, el tipo de usuarios que los utiliza, sus finalidades entre otros.

4.1.1 Definición y características

Tradicionalmente los metadatos se han definido como datos acerca de datos, pero esta definición no da a entender cómo es representada la información en su conjunto.

Para los estudiosos en diferentes áreas del conocimiento la palabra metadato puede significar diferentes cosas pero, en este trabajo la intención es conocer el significado de esta palabra para los bibliotecarios, de tal manera que los metadatos son definidos por varios autores del siguiente modo:

- ❖ Los metadatos son definidos como aquellos que proporcionan información acerca de documentación o de otros datos manejados con una aplicación o desarrollo. Los metadatos pueden incluir descripción de información acerca de su contexto, calidad y condición o características de los datos. (Taylor, 2004).
- ❖ Los metadatos son datos asociados con objetos (documentos como objetos) que liberan a sus usuarios un conocimiento avanzado completo sobre su existencia o características... (Dempsey y Heery, (1997, p. 5).
- ❖ Los metadatos son información inteligible por ordenador sobre recursos Web u otras cosas (Berners, 1997).
- ❖ Los metadatos son un término usado normalmente para referirse a aquella información sobre objetos digitales (Goodchild, 1995).
- ❖ “Literalmente, metadatos significa datos sobre datos. Los metadatos son datos asociados a objetos o sistemas de información para fines de descripción, administración, uso, preservación, etc. Existen varios tipos de metadatos: descriptivos (que sirven para identificación y localización),

administrativos (creación, derechos, control de acceso, etc.) y estructurales que relacionan los objetos” (Barrueco, 2010).

- ❖ Los metadatos se entienden, como información adicional necesaria para que los datos sean útiles, es decir, que facilitan la gestión, el acceso o el análisis de datos, funciones que serán la base de las aplicaciones o usos que se le dan a la metainformación en el contexto de Internet.

Coincidiendo, se puede resumir que los metadatos están destinados a ordenar y describir la información contenida en un documento entendido como objeto, de tal forma que éstos son como reveladores de la descripción formal y el análisis del contenido de los documentos de información, para mejorar el acceso a ellos en la Red. Son estructuras de la organización de la información, legibles por máquina, para hacer útiles los datos de distintas formas, según las necesidades concretas de cada servicio de información digital y según la aplicación que se les otorgue.

Los metadatos son información estructurada que describe los atributos de grupos de información con el propósito de identificar, descubrir y algunas veces administrar los datos. Pueden incluir información descriptiva acerca del contexto, calidad y condición o característica de los datos.

Los metadatos incluyen no sólo información descriptiva tal como se encuentra tradicionalmente, sino también información necesaria para la administración, uso y preservación de la misma.

Los metadatos son estructuras de base para describir distintos objetos de información distribuidos en la web, de tal forma que la búsqueda con metadatos potencie la recuperación de la información digital.

Los metadatos son datos que sirven para describir en forma codificada grupos de datos a los que se podrían llamar “objetos informáticos”. Son descripciones

estructuradas de un objeto informático. Literalmente son “datos sobre datos”, incluyen datos asociados tanto con un sistema de información como con un objeto informático para fines descriptivos, administrativos, legales, técnicos, de uso y conservación.

Los metadatos permiten simplificar la búsqueda para clasificar y visualizar la información, de modo que sea más fácil de usar y gestionar efectivamente una colección. Sus funciones principales son: facilitar la identificación, ubicación, recuperación, manipulación y uso de recursos de información accesibles en línea.

Los sistemas de metadatos tienen tres niveles de complejidad a saber:

- Un simple formato en el que el metadato no es más que un dato estructurado encontrado en un recurso.
- El segundo nivel es un formato estructurado, éste incluye metadatos formales y elementos que han sido creados para uso en general y cuenta con una plantilla para la creación de metadatos, no requiere de nivel profesional para la descripción.

El tercer nivel de complejidad es el que se ocupan en bibliotecas, archivos, museos y conjuntos de información, en el que se interesan ciertos grupos como son los repositorios institucionales de información.

La creación y uso de metadatos es una parte importante en todas las estrategias operacionales de preservación digital. Los metadatos informan las partes importantes del objeto digital e indican su localización, son una forma especializada de administrar, éstos pueden ser usados como medio de explotar la información técnica que soporta la preservación de los objetos digitales, además apoyan y facilitan la retención a largo plazo de la información digital.

Para los estudiosos en las diferentes áreas del saber, los metadatos pueden tener significados diferentes pero para los bibliotecarios son un tipo de datos que se usan para describir el contenido, la estructura, la representación y el contexto de algún conjunto de datos específico siguiendo estándares ya establecidos.

Una característica de los metadatos es su capacidad de relación o el establecimiento de enlaces, indizando y clasificando las inmensas cantidades de información de diversos tipos, por ello se han hecho imprescindibles en la recuperación de información en línea.

4.1.2 Tipos de metadatos

Existen diferentes tipos de metadatos, ellos se aplicarán dependiendo de las necesidades que se tengan para almacenar la información, pues las finalidades para el manejo de este tipo de información son diversas.

Son múltiples las clasificaciones de los metadatos que se pueden hacer en base al origen, forma, funcionalidad, estadística de uso, nivel de estructuración de los datos y más, pero una de las clasificaciones más sencillas es una división de metadatos, según el ámbito de aplicación en que se utilizan. Tomando en cuenta la dimensión de metainformación según el dominio de aplicación se tienen metadatos:

- Para describir recursos de información Web asociados al concepto de biblioteca digital (por ejemplo el modelo propuesto por el Dublin Core).
- Metadatos que definen los elementos de los registros catalográficos en bibliotecas (como es el caso del formato MARC).

- Metadatos asociados a la información de museos y por tanto, una información eminentemente iconográfica.
- Metadatos para recursos geográficos o geoespaciales.
- Metadatos para describir recursos de información electrónica gubernamental.
- Metadatos para la descripción archivística como alternativa a la descripción tradicional realizada en archivos. Así también se reconoce la capacidad de los metadatos para proporcionar información descriptiva en lo procesos de creación de registros electrónicos.
- Metadatos específicos en el área de desarrollo de dispositivos inteligentes para el almacenamiento de datos (Méndez Rodríguez, p. 57, 2002).

Entre las diferentes clasificaciones que tienen los metadatos, se tiene una clasificación sobre las distintas fases por las que pasa un documento electrónico entendido como objeto, la división según la utilidad y/o función que presenta en esas diferentes etapas de lo que se denomina ciclo vital de un objeto. Es así que se pueden establecer cinco tipos de metadatos como son:

- ✓ **Metadatos administrativos.**- Son utilizados para la gestión y administración de los recursos digitales en red en el momento de su creación o introducción en el sistema. Éstos son útiles para comprobar el nivel de mantenimiento de los recursos Web -por ejemplo: la ubicación de la información, la institución en cuyo servidor se alberga, seguimiento y control de las distintas versiones- Así también son los relacionados con el funcionamiento de los sistemas o el comportamiento

de los metadatos, por ejemplo la información sobre el hardware o el software, digitalización de información y demás.

- ✓ **Metadatos descriptivos.**- Usados para describir o identificar información sobre recursos de información en la fase de organización. Para la representación o identificación de los recursos de información en la fase de organización de los objetos digitales.
- ✓ **Metadatos para la conservación.**- Destinados a gestionar la preservación de las fuentes de información con la intención de resolver los principales problemas de la información digital, esto es: la fragilidad y el carácter efímero de los objetos y el problema de la obsolescencia de los medios informáticos (hardware y software).
- ✓ **Metadatos técnicos.**- Generados para un sistema automatizado relativo al funcionamiento del propio sistema.
- ✓ **Metadatos de uso.**- Son los relacionados con el nivel y el tipo de uso de los recursos de información como el seguimiento de uso por los usuarios. Para preservar fuentes de información como los documentos sobre renovación y actualización de datos o la migración del contenido a otro sitio (Méndez Rodríguez, p. 57, 2002).

Así, en cuanto a las funciones de los metadatos se tienen:

- Facilitar la interoperabilidad de sistemas para el intercambio de información
- Recuperación de la información
- Brindar puntos de acceso (recuperación)

- Codificar la descripción para ser recuperada automáticamente por los descriptores (localización y acceso)(Méndez Rodríguez, p. 48-49, 2002 y Angelozzi, p. 25-26, 2010).

4.1.3 Aplicación de los metadatos

Los metadatos tienen diferentes niveles de aplicación, por ejemplo:

- Para recursos de información local, estatal o federal, como las publicaciones gubernamentales o el sitio Web de una agencia federal o de una administración pública.
- Información museística como exposiciones de patrimonio cultural compuestas de un catálogo de exposición, imágenes, textos murales y arte folclórico.
- Libros, revistas, periódicos o colecciones de materiales relacionados pertenecientes a bibliotecas digitales (Méndez Rodríguez, p. 35, 2002).

Los metadatos son utilizados para recuperar diferentes formatos de información electrónica, por ejemplo:

- a) Navegación. Los metadatos en este caso servirán en como soporte para navegar y recuperar colecciones de imágenes que requieren un conocimiento sobre el contenido espacial de las mismas y de su comportamiento.
- b) La recuperación de imágenes en movimiento extraídas de un vídeo, cuya búsqueda deberá basarse también en el contenido semántico, captando el contenido espacial y también el aspecto temporal.

- c) Igualmente, para los registros de audio (por ejemplo sistemas de información de radio) de tal manera que se capture en forma textual, la esencia del contenido sonoro.
- d) La utilidad de los metadatos se refleja también en la aplicación de los mismos para la gestión de documentos estructurados como SGML y XML para la recuperación eficiente que se consigue a través de la explotación de esa estructura y de la utilización de los metadatos para la indización.
- e) Los sistemas de información geográfica y medioambiental; para este tipo de sistemas de información, es necesaria la integración que será posible a través de información descriptiva dirigida tanto al sistema como al usuario final. Esto implica la descripción de metadatos y la cohesión de vocabularios utilizados por los distintos sistemas de información para interpretar las descripciones.
- f) Acceso a distintos tipos de información. En este caso deberá contarse con sentencias de búsqueda formuladas independientemente del tipo de medio, para ello es necesario utilizar descripciones de metadatos para detallar, de manera uniforme, la información de los distintos medios.
- g) Otra aplicación es la del control de los recursos digitales, tanto para controlar el acceso, como para calificar el contenido o para describir los derechos del documento (Méndez Rodríguez, p. 54, 2002).

Los metadatos aumentan la probabilidad de que los usuarios sean capaces de recuperar información apropiada para sus consultas y puedan evaluar su relevancia (utilidad y accesibilidad), pues proporcionan mayor accesibilidad para la búsqueda por múltiples colecciones o bien, crear colecciones virtuales a partir de materiales distribuidos entre varios depositarios incorporando elementos comunes tales como la Encoded Archival Description (EAD), el Text Encoding Initiative (TEI)

o el Dublin Core, que permiten al usuario moverse con facilidad entre las copias descriptivas del objeto informático, las versiones digitales del objeto informático y las versiones digitales del objeto mismo y, al mismo tiempo, llevar a cabo búsquedas, tanto a nivel objeto como de la colección dentro y a través de los sistemas de información.

Otras características que tienen los metadatos son:

- ❖ **Relación de contexto**, es decir cuando se mantienen objetos que están interrelacionados entre sí, y a la vez, asociados con individuos, lugares, movimientos y acontecimientos (por ejemplo en museos, archivos y bibliotecas), lo que es lo mismo, que un objeto de una colección, después de ser digitalizado, se separa tanto de la información con que está catalogado como de sus relaciones con otros objetos de la misma colección; en este sentido los metadatos se aplican para el mantenimiento de las relaciones, así como en el establecimiento de la autenticidad, la integridad estructural y la exhaustividad de los objetos informáticos.
- ❖ **Uso más amplio de la información**. Con los metadatos correctamente estructurados es posible facilitar un número infinito de estrategias para las búsquedas de información, presentar los resultados e incluso manipular los objetos informáticos sin comprometer su integridad.
- ❖ **Versiónes múltiples**. Se refiere al incremento de información digital a partir de una versión original. La producción múltiple y distinta de los objetos, en este caso, los metadatos deben ser adoptados para que conecten las distintas versiones y captar las similitudes y las diferencias entre cada una de ellas.
- ❖ **Cuestiones legales**. Con el uso de los metadatos se permite a los depositarios identificar los numerosos niveles de información sobre los

derechos y las restricciones que afectan el uso de un objeto informático y sus distintas versiones. Así también los metadatos documentan otros requisitos legales, o aquellos impuestos sobre los objetos por los donantes.

- ❖ **Conservación.** Con el uso de los metadatos es posible resistir los cambios de nuevos sistemas para conservar la información, ya que independientemente del sistema que se use ésta podrá consultarse (Baca, p. 7-8, 1998).

Existen también los metadatos de propósito general, éstos son:

Aquellos formatos dirigidos a cualquier tipo de información, O sea, los destinados a describir cualquier Objeto de información. Son los destinados a la representación y descripción de documentos electrónicos que independientemente de su temática y/o finalidad informativa proporcionan datos de diverso tipo sobre los recursos. Ejemplo de ello es Dublin Core, METS y más.

También se conocen los metadatos con propósito específico que están vinculados a un tipo de información digital o a un dominio temático o comunidad informativa en particular como EAD, TEI, entre otros.

4.1.4 Esquemas de metadatos

Anteriormente se mencionó que los metadatos son una herramienta que en los últimos años ha tenido fuerza para la recuperación de la información vía web, puesto que con ellos es posible mantener organizada la información contenida en ella. Su función se enfoca a dos puntos de vista: la del sistema, en ella se facilita la interoperabilidad y la posibilidad de compartir datos dentro de las diferentes herramientas de búsqueda y la del usuario, en el que permite la recuperación del

conocimiento al suplir dos necesidades básicas de los usuarios: el de acceso y la recuperación de la información.

El establecimiento de normas y procedimientos estandarizados es fundamental para hacer compatibles los sistemas y facilitar la difusión y el intercambio automatizado de datos e información a través de redes de comunicaciones especialmente en Internet, ya que son pautas o directrices aceptadas por la comunidad, destinadas a promover almacenamiento consistente de datos, de tal forma que se convierten en piezas indispensables para el intercambio de información eficiente y para una búsqueda y recuperación de la misma.

Los metadatos son tan específicos de una disciplina como su vocabulario, por ello existen tantos metadatos como disciplinas que pretenden dar servicios de información a través de la web; es por ello también que se han creado varios modelos o esquemas de metadatos a partir de la realidad de los lenguajes sobre los que se constituye la World Wide Web.

Independientemente del área del conocimiento para las que se utilicen los esquemas, formatos o modelos de metadatos, deben dar cuenta de las particularidades de los recursos de información de la Red para facilitar su identificación y posterior recuperación. De igual manera todos los esquemas de metadatos tienen constantes comunes: un número limitado de elementos, un nombre para cada elemento o etiqueta y un significado asociado al elemento (Taylor, 1999). Estos esquemas coinciden también en tener una semántica descriptiva con información relativa al contenido, ubicación, atributos físicos tipo de objeto (texto, imagen, mapa, etc.) y forma.

Los elementos principales en la mayoría de los modelos de metadatos cubren el acceso a los documentos publicados en la WWW, incluyendo elementos como el creador, el título, cuándo y dónde se ha publicado, las materias que trata, entre otros.

Aunque todos los modelos de metadatos cuentan con tres elementos obligatorios (título, autor e identificador), son amplios y complejos como su finalidad y aplicación.

Tradicionalmente se ha procurado organizar la información, no importa tiempo ni unidad de información como son las bibliotecas, las hemerotecas, los museos, ludotecas, entre muchas otras. La forma más conocida que representa el contenido de este tipo de unidades es el catálogo.

Los estudiosos en el área de la bibliotecología se han preocupado por la organización de la información tanto tradicional (como son los documentos impresos), como electrónica. Para este tipo de información electrónica se ha encontrado que puede ser codificada asignando etiquetas, números etc.

Los proyectos que existen actualmente sobre metadatos no son normas sino proyectos de normas o estándares que se usan en determinados organismos y grupos de usuarios, pues no existen normas únicas internacionales.

Los metadatos se han constituido como una de las actividades que se han propuesto organizar los recursos informativos digitales, ayudando a identificarlos, describirlos y localizados.

Desde hace años los modelos de metadatos para la descripción de contenidos de los documentos, han proliferado crecientemente tratando de generar herramientas que permitan sustituir registros tradicionales para adaptarlos a objetos electrónicos, entre ellos se tienen: APPM, TEI, Dublin Core, GILS, FGDC, VRA, EAD y más.

Resulta difícil resumir todas las acciones, implicaciones y evolución de las iniciativas de metadatos del Dublin Core, ya que es el esquema más divulgado y más aceptado para describir recursos de información Web asociados con el

concepto de biblioteca digital; además de ser compatible con las Reglas de Catalogación Anglo Americanas y el ISBD (International Standard Bibliographic Description). No obstante, es el estándar más extendido para la representación y recuperación de la información Web y el que más influencia ha ejercido en lo que respecta al desarrollo general del concepto y teoría de los metadatos. Es el mejor, puesto que se adapta de acuerdo a la evolución de la World Wide Web.

Es así que Dublin Core desde su aparición en 1995, ha evolucionado para representar la información electrónica a través de 15 etiquetas de información básica. Los metadatos se sitúan de esta forma entre los formatos de descripción bibliográfica más tradicionales que se han visto obligados a contemplar en su normalización los recursos electrónicos como testimonian las ISB-ER de 1997, y los sistemas automáticos de organización y recuperación de los contenidos informativos.

4.2 The Dublin Core

El esquema de datos Dublin Core es el más utilizado en los repositorios institucionales que se detallan en OpenDOAR, Por ello, para propósitos de este trabajo será necesario contemplar el uso de **Dublin Core Metadata Initiative** conocido sólo como **Dublín Core**.

Originalmente Dublin Core fue diseñado para ayudar a los motores de búsqueda a encontrar y recuperar páginas Web y para la descripción de recursos, pero se ha convertido en el centro de atención de comunidades involucradas en la descripción formal de recursos en bibliotecas, museos, archivos, agencias gubernamentales, organizaciones comerciales, entre otras.

Dublin Core ha evolucionado hacia un formato de registro para el intercambio y recuperación de información del espacio digital.

El propósito general de Dublin Core es la recuperación de recursos interdisciplinarios de la Red a través de las siguientes actividades:

- 1) Desarrollar estándares de metadatos para la recuperación de la información en Internet a través de diferentes dominios.
- 2) Definir el marco para la interoperabilidad entre conjuntos de metadatos.
- 3) Facilitar el desarrollo de conjuntos de metadatos específicos de una disciplina o comunidad que trabaja dentro de la recuperación de información.

Algunas características que tiene Dublin Core para ser aplicable a la información digital son:

- Los elementos que lo constituyen son fáciles de entender y sencillos de usar.
- El conjunto de elementos es multidisciplinario e internacional en cuanto a su alcance y aplicabilidad
- Permite agregar modificaciones aplicadas a una disciplina o trabajo específico.
- Está en constantes revisiones y cambios, permite incluso, que diferentes áreas de información traten de adaptar este formato a sus necesidades.
- Todos los elementos son opcionales y repetibles, esto permite flexibilidad en la aplicación del formato y es posible establecer distintos niveles de detalle de la descripción.

- Soporta las cuatro habilidades que un usuario puede tener para recuperar la información: búsqueda, identificación, selección y recuperación.
- Debido a su cobertura internacional y ser multidisciplinario proporciona una base para la interoperabilidad semántica entre dominios de metadatos más ricos.

Dublin Core es implementado a través de HTML y puede ser codificado con XML y estructurado con RDF.

Dublin Core consiste en quince elementos que pueden ser divididos en tres grupos (Este rubro se tomó del libro de Taylor the organization of information como considerarse como fundamento):

- **Elementos relacionados al contenido del recurso**

- ✓ **Título:** El nombre del recurso (conjunto de información)
- ✓ **Materia y palabras clave:** El (los) tema(s) de la fuente (uso de vocabularios controlados y esquemas formales de clasificación ya establecidos)
- ✓ **Descripción:** Una descripción textual del contenido del recurso (podría ser un resumen o una descripción del contenido)
- ✓ **Fuente:** Información acerca de un recurso original del cual es derivado (aunque es recomendado usar una liga para dar la relación del elemento)

- ✓ **Idioma:** Indicación del idioma del contenido intelectual (texto) del recurso
 - ✓ **Relación:** El identificador de un segundo recurso, relacionado al recurso presente mediante una liga que incluye: Es versión de (IsVersionOf), es basado en (IsBasedOn), es parte de (IsPartOf), es formato de (IsFormatOf), etc.
 - ✓ **Cobertura:** Una identificación espacial (por ejemplo la región física) y/o temporal como fecha probable del contenido del recurso además una nota para el tema y abajo la fecha para la creación del recurso). La características espaciales y temporales deben ser tomadas de una lista controlada
- **Elementos relacionados al recurso cuando se ha visto la propiedad intelectual**
 - ✓ **Autor o creador:** El nombre de la persona u organización primaria responsable de la creación intelectual del contenido del recurso (por ejemplo: autor, artista, compositor, etc.)
 - ✓ **Editor:** El nombre de la entidad responsable para hacer el recurso disponible (como: casa editorial, universidad, algún departamento, otros cuerpos corporativos, etc.)
 - ✓ **Otro contribuidor:** El nombre de una persona u organización que ha hecho contribuciones significantes, pero aquellos que son secundarios al nombre del creador del elemento (por ejemplo: editor, ilustrador, etc.)

- ✓ **Manejo de derechos:** Una declaración, liga o identificador que proporciona información acerca del manejo de los derechos (como: tiempo de uso del recurso, restricciones de uso, tiempo en el que el recurso será removido del despliegue, etc.).

- **Elementos relacionados principalmente para el número del recurso como una instancia**
 - ✓ **Fecha:** La fecha de creación o disponibilidad del recurso; esto es recomendado por ISO 8601 (fecha y tiempo de los formatos) a ser usado, especialmente año-mes y día, como esquema de una fecha

 - ✓ **Tipo de recurso:** Una designación del tipo o categoría del recurso (como casa, poema, página, reporte técnico, diccionario, etc.). Es recomendable que se haga uso de una lista estandarizada.

 - ✓ **Formato:** Una designación del software y quizá el hardware requerido para usar el recurso. Es recomendable hacer uso de una lista estandarizada

 - ✓ **Identificador del recurso:** Una liga o número que únicamente identifique el recurso (como: URL, ISBN, etc.)

Los elementos como: tema, palabras clave, autor o creador, y otros contribuidores son elementales puntos de acceso.

Es importante señalar que Dublin Core tiene cuatro principios generales que son:

- 1) Puede ser expandido con otros elementos que se necesiten por una comunidad particular
- 2) Todos los elementos son opcionales
- 3) Todos los elementos son respetables
- 4) Los elementos pueden ser modificados. La forma del contenido de cada elemento no es prescrita.

4.3 Software

En los últimos años se han desarrollado varios software libres y comerciales que permiten crear y administrar repositorios de cualquier tipo con facilidad.

Es recomendable observar ciertas características en la plataforma de software que se utilice en el Repositorio Institucional, por ejemplo, deberá observarse que se puedan leer todo tipo de archivos como textos, imágenes, conjuntos de datos, vídeo, audio, etc., dado que necesario recordar que el repositorio contendrá diferentes tipos de materiales.

Existen numerosas plataformas para la creación y gestión de repositorios. Muchas de ellas son gratuitas y tienen licencias de uso de software libre como: DSpace, Eprints, Fedora, entre otros; otros son gratuitos pero no tienen licencias open source (como Zentity de Microsoft) y otras son plataformas comerciales como Digital Commons, Digitool, CONTENTdm, Digibib Open Repository, etcétera. La mayoría de plataformas para repositorios tienen funcionalidades básicas comunes, aunque con otras muchas específicas y características propias.

A pesar de que se ha recomendado el software libre, también se puede hacer uso de software propietario sólo que en este caso, depende de la empresa que lo

proporciona en todos los aspectos técnicos ya que sugiere gastos; en cambio, el software libre -que no quiere decir que sea gratuito, sino que se brinde un conjunto de libertades como ejecutar, copiar, distribuir, estudiar, cambiar y mejorar el software para adaptarlo a las necesidades que se tengan en la creación y manutención, en este caso, de un repositorio institucional- también depende de tecnologías libres como: tecnología java, servidores Apache, bases de datos MySQL, entre otros y se beneficia de la aportación de múltiples programadores de todo el mundo que mejoran el sistema continuamente o lo amplían con nuevas funcionalidades; en este caso, quizá algunos inconvenientes cuando se elige software libre son: que la institución debe ser capaz de instalar, adaptar, gestionar y mantener en funcionamiento el sistema por sí misma, a causa de que no se ofrece un servicio técnico y de apoyo por parte de los desarrolladores.

Para la creación de repositorios institucionales se manejan aquellos paquetes de software dirigidos a la implementación integral de repositorios que cumplan con los estándares provistos por OAI-PMH. Generalmente estos paquetes no necesitan excesiva configuración o adaptación por parte de los administradores. Estos sistemas ya instalados y configurados en un servidor web deben ofrecer una serie de funcionalidades básicas para el mantenimiento y gestión del repositorio como son: la gestión de soporte de flujo documental (remisión de documentos; evaluación, aceptación o rechazo de los documentos), la edición, evaluación y transformación de metadatos sobre los recursos, interfaces de usuario de autor y de administrador, permiten la creación de diversos grupos de usuario y niveles de acceso; en cuanto al contenido, facilitan la creación de colecciones de materiales y generalmente aceptan múltiples formatos de archivo; para el uso del repositorio se limitan a diversas opciones de búsqueda y navegación entre los registros del mismo (palabras clave, búsquedas por campos de metadatos), así como para su visualización y descarga.

Cabe mencionar que existe un gran número de software que pueden aplicarse a la creación de repositorios institucionales como son: DSpace, Fedora, Architecture, Greenstone, GNU EPrints, CDS Invenio, OPUS, dLibra, Drupal, Digital Commons, por mencionar algunos. Sin embargo, en este apartado sólo se tratarán DSpace, EPrints, y Greenstone que por sus características, grado de implementación, estabilidad y su gratuidad pueden adaptarse en la creación de repositorios institucionales.

4.3.1 DSpace

DSpace (<http://www.dspace.org>) ha sido desarrollado conjuntamente por los laboratorios Hewlett-Packard y por la biblioteca del MIT (Massachusetts Institute of Technology) específicamente para constituir servicios de repositorio institucional.

DSpace es el software más empleado en los repositorios de acceso abierto que describe el Directorio de repositorios de acceso abierto (OpenDOAR), por ello se considera el más reconocido en la implementación de éstos.

DSpace es un software de código abierto que proporciona organización para capturar y describir material digital usando módulos de envío con una variedad de opciones. El sistema permite distribuir informaciones vía web de manera que puedan ser recuperadas mediante búsquedas, así como almacenar y preservar objetos digitales por largos periodos de tiempo. Soporta una gran variedad de datos incluyendo libros, tesis, fotografías, vídeo, datos de investigación y otras formas de contenido.

Este software permite:

- Recoger y describir material digital usando un módulo de presentación del flujo de trabajo.

- Distribuir los recursos digitales de una organización sobre la web a través de un sistema de búsqueda y recuperación.
- Preservar contenidos digitales a largo plazo.

La forma en que DSpace está construido permite una distribución jerárquica de tareas entre un grupo de personas, por tal razón, es posible tener repositorios multidisciplinarios organizados temáticamente y manejados por administradores delegados. Es una plataforma diseñada para la captura, almacenamiento, indización, preservación y redistribución de la producción intelectual del personal investigador de la universidad en formato digital distribuido mediante una licencia BSD (*Berkeley Software Distribution*).

DSpace está construido en base al concepto de comunidades y subcomunidades de una institución y sus colecciones. Las comunidades representan unidades organizativas, como departamentos académicos o administrativos, centros de investigación, laboratorios e incluso investigadores individuales. Además DSpace organiza los usuarios en cuentas personales (*e-person*) y grupos, obligando a todos los usuarios personales a pertenecer a un grupo definido para poder asignarle permisos de escritura, lectura, modificación y eliminación de registros. Estas autorizaciones pueden realizarse a nivel de comunidad o colección.

Cada colección está compuesta por ítems, en donde son los elementos básicos del archivo. Cada ítem es propiedad de una colección. Cada ítem puede aparecer también otras colecciones, sin embargo solo puede pertenecer a una colección.

Los ítems también pueden estar subdivididos en paquetes de bitstreams. Estos bitstreams son flujos de bits, normalmente archivos. Los bitstreams que de alguna manera están relacionados, como por ejemplo los archivos HTML y las imágenes que componen un documento HTML, están organizados en paquetes.

DSpace está programado en Java y corre sobre sistemas tipo Unix, aunque ya se han realizado con éxito pruebas de instalación en Windows XP. Precisa de un servidor web Apache y una base de datos relacional de código abierto PostgreSQL 8.0 o superior, permitiendo también el uso de Oracle 9.0 o posterior. Aunque soporta todo tipo de formatos de documentos, recomienda a los administradores de cada sistema decidir qué formatos aceptar con fines de preservación, pudiendo diferenciar así tres categorías de formatos: soportados, conocidos y no soportados.

Otras característica técnica de DSpace es el empleo de identificadores persistentes de recursos mediante el sistema Handle del CNRI (Corporation for National Research Initiatives). Además, el interfaz web básico que ofrece DSpace por defecto es totalmente configurable según las necesidades de cada institución o comunidad de usuarios, mediante HTML y Java.

DSpace utiliza el motor de búsqueda de software libre Lucene, permitiendo búsquedas tanto por campos de metadatos como en el texto completo, así como realizar búsquedas booleanas.

Dentro de DSpace los metadatos juegan un papel importante, porque en base a éstos se agregan los ítems dentro del repositorio, por defecto DSpace maneja el esquema de metadatos Dublin Core (dc), este esquema es muy general ya que maneja elementos como: título, fecha, autor, etc.

DSpace tiene la opción de agregar metadatos al esquema Dublin Core o en su defecto, agregar nuevos esquemas, siempre y cuando se siga el estándar Dublin Core.

DSpace sistema es desarrollado para constituir servicios de repositorio institucional; es diseñado para la captura, almacenamiento, indización, preservación y distribución de la producción intelectual del personal investigador

en formato digital. Este sistema está construido en departamentos académicos o administrativos, centros de investigación, laboratorios, e incluso investigadores individuales. Este software organiza a los usuarios en cuentas personales (e-person) –porque algunos usuarios pueden ser máquinas en lugar de personas- y grupos, para que los usuarios individuales pertenezcan a un grupo definido para poder otorgarle permisos de escritura, lectura, modificación y eliminación de registros.

DSpace está programado en Java y corre en sistemas tipo Unix, necesita un servidor web Apache, una base de datos de código abierto como PostgreSQL 8.0 o superior y Oracle 9.0 o posterior. Aunque soporta todo tipo de formatos de documentos, es recomendable decidir el tipo de formatos que se aceptarán con fines de preservación.

La interfaz de DSpace es totalmente configurable según lo necesite cada institución mediante HTML y Java. Además utiliza un motor de búsqueda de software libre Lucene, permitiendo la búsqueda tanto por campos de metadatos como en texto completo, así como realizar búsquedas booleanas.

Características de DSpace:

- ❖ Es posible mantener una organización jerárquica de contenidos.
- ❖ Se adapta a las necesidades del usuario en aspectos como interface de administrador y usuario final.
- ❖ Se instala fuera de línea.
- ❖ Usa código basado en estándares.

- ❖ Usa un motor de bases de datos PostgreSQL, lo cual permite corregir problemas fácilmente.
- ❖ Permite a los autores poner sus trabajos inmediatamente después de su creación a disposición de los usuarios.
- ❖ Es posible dar seguimiento a los trabajos mediante referencias duraderas en tiempo, ofreciendo estadísticas de uso.

4.3.2 EPrints

GNU EPrints (antes EPrints) (<http://www.eprints.org>) ha sido desarrollado en la Universidad de Southampton y actualmente está financiado por JISC (Joint Information Systems Committee) del Reino Unido, como parte del Open Citation Project, y por la NSF (National Science Foundation) de EEUU. En principio está diseñado para la creación de repositorios de e-prints, aunque puede ser empleado con otros objetivos.

Hasta el momento, el software GNU EPrints es el segundo sistema más utilizado en la creación de repositorios (DSpace le ha superado recientemente), El gran número de instalaciones de GNU EPrints sugiere que una institución puede instalarlo y ponerlo en marcha de forma relativamente rápida, y con un mínimo de habilidad técnica.

EPrints soporta artículos que pueden ser electrónicos, libros, capítulos de libros, monografías, informes, conferencias, tesis, patentes, entre otros formatos.

EPrints usa formatos de texto como: JPEG, PNG, GIF, BMP, formatos de archivo TIFF, formatos AVI, MPEG y formatos MS, PowerPoint, MSWord y archivos zip3Tar.gz y es compatible para la integración de objetos digitales y el soporte con

la importación de los metadatos de los objetos de DOI, XML, Pubmed ID, formato PubMed XML, XM2/PubMed, además es fácil importar metadatos si está disponible en XML.

EPrints además es también compatible con: Cita ASCII, BibTex, Dublin Core, DID2, ED3, con archivos incrustados en: EndNote, HTML Cita, METS, MOODS, Open URL y con objetos de contenido Open URL DisertTION, Diario Open URL, Consulte, Reference Manager y metadatos simples.

Actualmente en su versión 3.0, el código de E-prints está escrito en Perl, utiliza el servidor web Apache y MySQL como base de datos, pudiendo funcionar sobre los sistemas operativos Linux, GNU o Solaris, así como en Windows gracias a un módulo adicional desarrollado su última versión.

Aunque por defecto se emplean metadatos Dublin Core, EPrints permite aplicar cualquier esquema de metadatos, y que el administrador decida qué campos se asignan a cada tipología documental.

EPrints permite la recuperación de contenido mediante búsqueda textual (simple o avanzada, variando el número de campos de metadatos en los que se puede consultar) y navegación. En cuanto a formato, un archivo EPrints puede almacenar cualquier tipo de ficheros que el administrador decida aceptar, incluyendo HTML, Adobe PDF, y PostScript, y cada documento puede ser almacenado en varios formatos distintos.

En EPrints el proceso de remisión de contenido contempla las tareas de evaluación y revisión de metadatos, comprobación y conversión de formatos, y las opciones de aprobación, rechazo o reenvío del documento al autor para su modificación.

GNU Eprints permite crear identificadores persistentes, aunque se recomienda asignar identificadores propios independientemente del sistema. En general, se

considera que este sistema funciona mejor para documentación textual y no multimedia, y no soporta formatos para objetos complejos como METS, MPEG21 o IMS CP (para el empaquetamiento de objetos de aprendizaje), por lo que no se suele emplear por bibliotecas digitales, repositorios de objetos de aprendizaje, ni por aquellos repositorios que pretendan albergar documentación gráfica, sonora o audiovisual.

EPrints es una plataforma flexible que brinda libertad para ampliar su funcionalidad, puede manipular una gran cantidad de objetos digitales, desde objetos textuales hasta objetos multimedia. No está limitado comunidades y colecciones estructuradas, permite la creación de colecciones virtuales flexibles a partir de los metadatos. Crea un identificador numérico para cada documento que se añade al repositorio y Alerta automáticamente sobre el ingreso de posibles depósitos duplicados. Esta plataforma está diseñada para asistir al auto-archivo.

La instalación de EPrints puede bajar múltiples repositorios y registra automáticamente todos los accesos a los registros bibliográficos y a los textos completos.

4.3.3 Greenstone

La plataforma Greenstone aunque no ocupa un lugar primordial de uso en OpenDOAR, cuenta con bondades que optimizan el funcionamiento en la implementación de repositorios académicos.

Greenstone tiene por objetivo explorar el potencial de bibliotecas digitales accesibles por Internet y desarrollar sistemas que impongan una estructura a repositorios de información que carecen de catalogación y distribución.

Greenstone es un conjunto de programas de software diseñado para la construcción y distribución de colecciones digitales proporcionando así una nueva forma de organizar y publicar la información a través de internet o en forma de CD-ROM.

Como objetivo Greenstone tiene: brindar el potencial de construir sus propias bibliotecas digitales a los usuarios con repositorios documentales proporcionando una contribución importante al acceso universal a la información, un apoyo a la libertad de información y a la democratización del conocimiento. Es un paquete multiplataforma fácil de utilizar multilingüe que permite el montaje de documentos electrónicos en las colecciones digitales y su publicación en la web. Acepta documentos en una amplia gama de formatos y es compatible con numerosas normas de intercambio de documentos y metadatos, incluyendo el cumplimiento del protocolo OAI-PMH.

Greenstone es de fácil instalación y se aseguran las mínimas exigencias de recursos del sistema siendo de fácil instalación, en esta plataforma es posible la compatibilidad con versiones anteriores.

La flexibilidad y facilidad de uso y libre disponibilidad de Greenstone lo convierten en un recurso especialmente útil para los países en desarrollo.

Las aplicaciones de Greenstone son múltiples, una ventaja central de esta plataforma es su disponibilidad y funcionamiento en diferentes lenguas. La definición de metadatos en diferentes idiomas permite la recuperación de archivos en estos mismos idiomas.

Greenstone se soporta en Windows y Unix, además puede desarrollarse sobre servidores Web Apache, PWS o IIS de Microsoft y requiere de aplicación Java para su funcionamiento. Debe contar también con algún mecanismo que permita almacenar y/o exportar los objetos digitales no importando el formato utilizando el

lenguaje XML. Las computadoras clientes pueden mostrar las colecciones en ambiente Web. No necesita ningún tipo de configuración por parte de sus usuarios.

Greenstone es interoperable usando estándares actuales y soporta la utilización de esquemas de metadatos estándar. El usuario es quien define el esquema a utilizar, este conjunto de metadatos ya vienen predefinidos como: Dublin Core, CDS/ISIS, ProCite, BibTex, Refer, OAI, DSpace y METS.

Los formatos de documentos textuales soportados son: PDF, Post Script, Word, RTF, HTML, TXT, Latex, Excel, PPT, Email, código fuente, Open Office. Para documentos multimedia: imágenes en cualquier formato incluyendo: GIF, JIF, JPEG, TIFF, archivos de audio MP3, Ogg Vorbis, MIDI, etc., también es capaz de procesar archivos zip, descomprimiéndolos al momento de importarlos.

En el ambiente de trabajo de Greenstone, se asignan metadatos a los documentos mediante una plantilla de acuerdo al modelo de metadatos escogido aunque el protocolo funciona de manera nativa con metadatos Dublin Core y un formato propio que ofrece esta plataforma. Es posible asignar metadatos a los documentos individuales o a varios, agrupados en carpetas. La página principal de cada colección, tiene una breve descripción del tema de la colección y las formas en las que pueden realizarse las búsquedas. La recuperación puede hacerse por medio de los índices que asigne el creador de la colección; incluye autor, título, materia y el texto completo; además puede navegarse en la colección mediante la consulta de listas también asignadas por el creador de la colección, principalmente autor, temática y fecha.

En base a lo anterior, se puede concluir lo siguiente:

- ❖ DSpace es la solución más adecuada cuando se necesita implementar un repositorio institucional que dará soporte a diferentes tipos de documentos y atender a varias comunidades de usuario, gracias a su versatilidad.
- ❖ Greenstone es un software para crear bibliotecas digitales en un entorno de baja intromisión externa. No existe acuerdo sobre la conveniencia de usar este software en la implementación de repositorios institucionales.
- ❖ EPrints es la aplicación correcta cuando se necesita implementar una colección de pre-prints o revistas digitales (Tramullas y Garrido Picazo, p. 176, 2006).

Conclusión

Por el hecho de que la mayor parte de repositorios institucionales analizados en OpenDOAR hacen uso del software DSpace, es evidente que se hace uso del esquema de metadatos Dublin Core en la implementación de repositorios institucionales.

En este trabajo se tomaron para su descripción sólo tres software que son: DSpace, EPrints y Greenstone, ya que de las muchas plataformas de acceso abierto para colecciones digitales son los más destacados. DSpace y EPrints son los mayormente adecuados para repositorios institucionales, en cambio Greenstone es el mayormente establecido para la administración de colecciones y bibliotecas digitales.

EPrints hasta el momento, es el segundo sistema más utilizado en la creación de repositorios de libre acceso. Emplea cualquier esquema de metadatos.

Es un software de libre distribución y sujeto a la licencia Pública General (GNU), esto significa que su código fuente es abierto y libremente modificable por cualquier programador que quiera realizar un cambio pero con la condición de que las modificaciones sean gratuitas y en abierto. Es posible un ajuste para cada entorno, incluso diferente a lo recomendado, pero esto puede relacionarse con un gasto considerable.

En general se considera que esta plataforma funciona mejor con documentación textual y no multimedia no soporta formatos complejos como METS, MPEG21 o IMS CP (para el empaquetamiento de objetos de aprendizaje) por lo que no se suele emplear en bibliotecas digitales, repositorios de objetos de aprendizaje ni por aquellos repositorios que pretendan albergar documentación gráfica, sonora o audiovisual.

Teniendo en mente lo descrito, la plataforma de software EPrints se limita a guardar la producción intelectual en formatos textuales; esto limita el objetivo de un repositorio –guardar toda la información institucional creada por sus investigadores, no importando el formato o tipo de documento-, es por ello que se descarta su aplicación para repositorios institucionales.

Dspace permite que su esquema de metadatos pueda ser “combinado” con otros esquemas para una adecuada gestión de los documentos. Otro elemento que permite la ágil recuperación de la información, son los identificadores digitales; una serie de secuencias cortas de caracteres alfanuméricos que identifican recursos en la red y los hacen disponibles con una amplia variedad de protocolos. Es un sistema que provee la forma de gestionar materiales y publicaciones resultantes de las diversas investigaciones por parte de los investigadores, almacenadas en un repositorio que ofrece a los usuarios accesibilidad en todo momento. Utiliza el estándar Dublin Core para la descripción de los documentos que van desde texto a lo más sofisticado y posteriormente posibilita su recuperación precisa. Los usuarios pueden hacer sus búsquedas a partir de los metadatos declarados por

los especialistas o simplemente mediante las listas de documentos por campos específicos, como son: autor, título y fecha.

EPrints y DSpace se basan en diferentes tecnologías pero son casi idénticos en su funcionalidad (funciones de búsqueda, archivo de documentos, interfaces en línea para el autoarchivo...).

En cambio Greenstone es un conjunto de programas y aplicaciones de software diseñados para la creación y difusión de colecciones digitales. Greenstone acepta todo tipo de formatos como son: texto, imágenes, bases de datos, etc., y su procesamiento se realiza fácilmente. En este ambiente de trabajo se asignan metadatos a los documentos mediante una plantilla, según el modelo de metadatos escogido; entre los que está Dublin Core y un formato que ofrece Greenstone.

DSpace y Greenstone son sistemas que tienen comunidades de código abierto, pero aún así existen algunas diferencias entre ambos software planteados:

Preservación:

DSpace se orienta exclusivamente hacia la conservación a largo plazo e incluye un esquema donde las instituciones se comprometen a garantizar la continua disponibilidad.

Greenstone es para la organización y creación de colecciones documentales digitales para publicarlas en Internet o en medios extraíbles como el CD-ROM.

Infraestructura de apoyo:

DSpace está diseñado para uso institucional donde hay instalaciones informáticas centralizadas y una infraestructura competente para soporte de software.

Greenstone está diseñado para facilitarlo a cualquier persona con conocimientos básicos en informática.

Operación multiplataforma:

DSpace se limita actualmente a Unix y OS/X, aunque esto no es una desventaja.

Greenstone funciona en ordenadores Windows, Unix y Mac OS/X.

Autor orientada:

DSpace incorpora una interfaz mediante el cual los autores generalmente pueden presentar documentos al sistema y definir sus metadatos.

Greenstone, no lo hace.

Bibliotecario orientada:

DSpace ofrece un diseño genérico que se puede adaptar, pero no por los usuarios finales.

Greenstone suministra una interfaz de usuario final con el que las colecciones se pueden diseñar, personalizar y construirlas.

Construcción de estándar de metadatos:

DSpace impone una única norma de metadatos en todas las colecciones.

Greenstone ofrece un estándar ampliamente utilizado (Dublin Core) pero también permite recolección de constructores a utilizar su propio esquema de metadatos,

ya sea mediante uno ya existente o mediante uno completamente nuevo con un editor de conjunto de metadatos.

Distribución de medios extraíbles.

En Greenstone es posible crear colecciones y pasarlas a una unidad extraíble como el CD-ROM de instalación automática que funciona en todos los sistemas Windows.

Colecciones dinámicas:

DSpace opera de forma incremental aunque las operaciones relativas al documento recientemente añadido como texto extraigan las miniaturas de las imágenes que se procesan.

Greenstone añadir documentos a una colección normalmente implica la reconstrucción de los índices y de navegación de estructuras de texto completo.

Usuarios internacionales:

Greenstone proporciona interfaces para los lectores en 35 idiomas y tiene un esquema que ayuda a que el lenguaje de las interfaces estén al día cuando se agreguen nuevas características de la interfaz.

Se debe tomar en cuenta que las características mencionadas, pueden cambiar a futuro.

Las diferencias en ambas plataformas se explican por los entornos para los que están hechos para operar. Al fin de cuentas ambos están diseñados para ayudar a terceros; sin embargo, tienen perspectivas diferentes y en muchos aspectos son complementarios. La característica que comparten es ser flexibles; ambos se

pueden personalizar y modificar en muchos niveles diferentes incluyendo el nivel de programación, ya que los dos son sistemas de código abierto.

Conclusiones

La democratización del conocimiento es un tema central en el proceso de comunicación científica, en medida que ha sido relevante superar las restricciones comerciales que aún rigen las diferentes casas editoriales en términos de comercializar el conocimiento científico.

Los mecanismos de colaboración entre todas las comunidades académica ha dotado de tácticas flexibles de intercambio sobre resultados de de avances tecnológicos.

La inmensa variedad de información es un fenómeno a tomar en cuenta, por lo tanto, a través de las redes se hace difícil el acceso a la información relevante, por lo que es necesario establecer buenos canales de comunicación para dar a conocer la información especializada.

La necesidad de compartir información por parte de académicos e interesados en obtener conocimientos actualizados, ayudó en gran medida a manifestar las inconformidades y lograr el Movimiento por el acceso abierto apoyado principalmente por tres iniciativas: Declaración de Budapest, Declaración de Bethesda y Declaración de Berlín.

El acceso abierto es el término utilizado para describir el libre acceso a la literatura científica en línea, en el cual el autor es el actor principal y beneficiario de éste.

La comunicación entre científicos modificó los mecanismos tradicionales de publicación y comunicación científica en cuanto a la divulgación de las tres Declaraciones en favor del acceso abierto que cuenta con numerosas ventajas.

En años pasados se comenzó a crear información y el compromiso de preservarla y darla a conocer. El uso de las TIC rebasó fronteras permitiendo la diseminación en todas las áreas del conocimiento así también, el uso de internet facilitó la difusión del conocimiento, lo cual se convirtió en un medio, por excelencia, para la difusión del conocimiento y como derivado, la creación de repositorios.

Los repositorios institucionales en acceso abierto son rentables para los investigadores en virtud de que son una vía inmediata para recibir resultados de investigaciones actualizados. Éstos cuentan con características como son: auto-archivo y la interoperabilidad, además de conservar la información a largo plazo y no existen estándares ni normas que los regule. Además es posible encontrar en ellos cualquier tipo de material.

Los repositorios institucionales son derivados del acceso abierto, éstos proporcionan una serie de servicios a los académicos quienes son los beneficiarios de ello, pues sus trabajos están mundialmente disponibles a través de los recolectores globales y sus ideas universalmente difundidas a través de los motores de búsqueda.

Los repositorios institucionales brindan un servicio que ofrecen las instituciones a los miembros de su comunidad y dan acceso a la información y garantizando la preservación de ésta a largo plazo.

Por lo anterior, esta nueva manera de conseguir información académica (acceso abierto) requiere el planteamiento de políticas y mecanismos de comunicación diferentes, en cuyos flujos de difusión sean más rápidos y efectivos.

Antes de implementar un repositorio es importante definir el esquema de metadatos que se aplicará.

Para la creación de repositorios institucionales, es recomendable hacer uso de un software libre y esquemas de metadatos de acceso abierto con el propósito de ahorrar recursos financieros, pues los metadatos son una herramienta importante para la organización de la información virtual vía WEB.

Es claro que entre las motivaciones que tienen los autores de artículos de revistas está el de comunicar los resultados de su investigación, ser leído y posteriormente citado, pero habría que preguntarse cuáles son las causas por las que muchos de los autores no hacen uso total de los repositorios, debido a que ellos no realizan el auto-archivo sus trabajos y en consecuencia no superan sus miedos o su apatía ante esta tecnología.

En esta situación de resistencia por parte de muchos autores ante el acceso abierto como son los repositorios, es recomendable tener en mente lo siguiente:

- Hacer publicidad del repositorio.
- Ofrecer el entrenamiento necesario.
- Ayudar a los autores a integrar sus artículos en el repositorio.
- Convencer a los autores de los beneficios del auto-archivo.
- Ayudar institucionalmente a otras dependencias.
- Someter los artículos a revistas o publicaciones dea.

Es recomendable también desarrollar una política de información y promoción de la filosofía del acceso abierto, pues una razón fundamental es el desconocimiento.

Las instituciones educativas y la sociedad han generado un sinnúmero de recursos informativos digitales en forma de texto, imágenes, vídeos, entre otros, que en ocasiones pasan inadvertidos y no son aprovechados debido a la carencia de metadatos que facilitan su ubicación, recuperación y gestión.

Es por ello que los aspectos tecnológicos, los metadatos y el hardware son los principales para la recuperación de la información.

El software libre ofrece estabilidad y flexibilidad en su manipulación, también permite interoperar y potenciar el uso de información independientemente del tipo de los materiales digitales existentes en una institución.

En los resultados obtenidos se observó que de 194 países con los que se cuenta actualmente a nivel mundial, sólo participaron 117.

Los continentes con mayor implementación de repositorios institucionales son: El Continente Europeo (1241) y en segundo lugar se encuentra Norte América (543).

Entre los países más desatacados por la implementación de repositorios digitales se encuentran: Estados Unidos con 455 (41%), Inglaterra 227 (21%), Alemania 169 (15%), Japón 145 (13%) y España 115 (10 %). El resto de los países de los diferentes continentes presenta una cantidad menor de repositorios.

En el estudio realizado mediante el Directorio de Repositorios de Libre Acceso (OpenDOAR), se observó que el software más utilizado es DSpace en los repositorios analizados, este software suma un 53 %, EPrints 18%, en cambio Greenstone suma un 3%. Para el desarrollo de este trabajo, es recomendable el uso de software libre y en este análisis se justifica el hecho de hacer uso de estos software libres y reconocidos a nivel mundial.

En la tipología, según OpenDOAR, se observa en los repositorios que la mayor parte la toman los artículos y los trabajos no publicados que suman el 22% cada uno de ellos, seguidos de las tesis (17%) y los trabajos no publicados (12%). En estos repositorios se tienen otro tipo de documentos como son: Grupos de datos, referencias, objetos de aprendizaje, software, libros, multimedia, trabajos

especiales, conferencias y patentes que realmente no figuran en la gráfica por ser tan pocos.

Aunque no hay información abundante relativa a los costos de implementación de los repositorios institucionales registrados en OpenDOAR, se considera que su costo es bajo, tal vez por la existencia de una gran cantidad de plataformas de software libre que permiten la gestión de documentos digitales. Sin embargo, en este aspecto es necesaria una inversión importante de recursos humanos, sobre todo en la etapa inicial de implementación, además de un servidor potente y de acceso rápido.

Como pudo observarse, la mayor parte de repositorios se centra en Estados Unidos y Europa, la plataforma de software más utilizados son DSpace e EPrints; Greenstone aunque no tenga relevancia en el estudio realizado, es conveniente decir que por sus bondades y debido a que es un software de código abierto, optimiza el funcionamiento de colecciones digitales que también se encuentran en repositorios.

Por el hecho de que la mayor parte de repositorios institucionales analizados en OpenDOAR hacen uso del software DSpace, es evidente que se hace uso del esquema de metadatos Dublin Core en la implementación de repositorios institucionales.

Entre los software DSpace, Eprints y Greenstone que si bien, son para aplicarlos en colecciones digitales, se encuentran diferencias relevantes como son:

- ✓ EPrints aunque sea el segundo software más utilizado internacionalmente, por lo general, se limita a guardar la producción intelectual en formatos de texto haciendo de lado los formatos gráficos, es por ello que se descarta su aplicación para que los repositorios institucionales cumplan con su objetivo principal.

- ✓ DSpace es más compatible y accesible para los usuarios en todo momento y permite la introducción de todo tipo de documentos en un repositorio, posibilitando su recuperación precisa.
- ✓ Greenston es un software diseñado para la creación y difusión de colecciones digitales.

DSpace y Greenstone son más establecidos y con trayectoria, pero cada uno se aplica a ambientes de trabajo diferentes; no quiere decir que DSpace no se pueda aplicar en bibliotecas digitales o que Greenstone no resulte en repositorios digitales, sino que cada software funcionará óptimamente para lo que está creado.

Admisible comentar que estos dos software mencionados anteriormente (DSpace y Greenstone), tienen un fin común: están diseñados para ayudar a terceros; sin embargo, tienen perspectivas diferentes y en muchos aspectos son complementarios. La característica que comparten es ser flexibles; ambos se pueden personalizar y modificar en muchos niveles diferentes incluyendo el nivel de programación, ya que los dos son sistemas de código abierto.

Por último, los resultados obtenidos relacionado a los contenidos que forman parte de los repositorios analizados a nivel mundial, indican que son usados académicamente.

Obras consultadas

Abadal, E. (2012). Acceso abierto a la ciencia.

Recuperado de:

<http://diposit.ub.edu/dspace/bitstream/2445/24542/11262142.pdf>

Abadal, E. (2012). Retos de las revistas en acceso abierto : calidad, cantidad y sostenibilidad económica.

Recuperado de:

<http://www.upf.edu/hipertextnet/numero-10/retos-revistas-en-acceso-abierto.html>

Agegn Alemu, G. (2009). The role of Open Access in fostering knowledge sharing and collaboration in Ethiopia : a case study. *Master thesis. International Master in Digital Library Learning*. Università Degli Study di Parma, Italia.

Recuperado de:

http://eprints.org/13108/1/Open_Access_Getaneh_Agegn_Alemu_Paper.pdf

Aguirre, S., Quemada, J. y Salvachua, J. (2013). Mediadores de interoperabilidad en e-learning.

Recuperado de: <http://jugla.dit.upm.es/~saguirre/publications/virtualEduca2004.pdf>

Angelozzi, S.M., Martín, S.G. (2010). Metadatos para la descripción de recursos electrónicos en línea: análisis y comparación. Buenos Aires : *Alfagrama Ediciones*. 147 p.

Antelman, K. (2004). Do Open Access articles have a greater research impact?.

Recuperado de:

http://eprints.rclis.org/bitstream/10760/5463/1/do_open_access_CRL.pdf

Arencibia, J. R. (2006). Las iniciativas para el Acceso Abierto a la información científica en el contexto de la web semántica. *Biblios*, Año 7 (25-26), 1-14.

Arriola Navarrete, O. (2011). Open Access y software libre : un área de oportunidad para las bibliotecas. *Biblioteca Universitaria*. Enero-junio 2011. 14 (1), 26-40.

Baca, M. (1999). Introducción a los metadatos : vías a la información digital. Washington. *Library f Congress Cataloging*. 43 p.

Bailey, C. W. (2005). Open Access Bibliography, liberating scholar literature with EPrints and Open Access Journals.

Recuperado de:

<http://www.digital-scholarship.org/cwb/WhatIsOA.htm>

Bainbridge, D., Osborn, W., Witten, I. H. y Nichols, D. M. (2006). Extending Greenstone for Institutional Repositories.

Recuperado de:

<http://www.cs.waikato.ac.nz/~iwh/papers/06-DB-et-al-ExtendingGreenstone.pdf>

Barrionuevo Almuzara. L. (2009). Open Access : la información científica al alcance de la sociedad.

Recuperado de:

<http://eprints.rclis.org/bitstream/10760/13098/1/comunicai%c3%b3n.pdf>

Barrueco J. M. y Subirats Coll, I. (2003). Open archives initiative. Protocol for metadata harvesting (OAI-PMH) : descripción, funciones y aplicaciones de un protocolo. *El profesional de la información*, 12 (2), 99-106.

Barrueco, J. M. (2003). OAI-PMH : protocolo para la transmisión de contenidos en internet.

Recuperado de:

<http://eprints.rclis.org/4093/1/cardedeu.pdf>

Barrueco, J.M., García Testal, C. (2009). Repositorios institucionales universitarios : evolución y perspectivas.

Recuperado de:

<http://bdigital.ucaldas.edu.co/bitstream/001/46/1/repositorios.pdf>

Barton, M. R. (2005). Creating an institutional repository : *LEADIRS Workbook*.

Recuperado de:

http://dspace.mit.edu/bitstream/handle/1721.1/26698/Barton_2004_Creating.pdf?sequence=1

Berlin declaration on open Access to knowledge in the sciences and humanities. (2003).

Recuperado de:

http://www.zim.mpg.de/openaccess-berlin/berlin_declaration.pdf

Berners-Lee (1997). Metadata architecture : document, metadata and links.

Recuperado de:

<http://www.3.org/DesignIssues/Metadata.html>

Bethesda statement on open access publishing (BSOAP, 2003).

Recuperado de:

<http://www.earlham.edu/peters/fos/Bethesda.htm>

Bories Fachin, G., R. Stumm, J., Lunardi Comarella, R., Fialho, F. A. P. y Santos, N. (2009). Gestión do conhecimento e a visão cogitiva dos repositórios institucionais. *Perspectivas em Ciência da informação*. 14 (2), 220-236.

Open Society Institute (2013). Budapest Open Access Initiative.

Recuperado de:

<http://www.soros.org/openaccess>

Bueno de la Fuente G. y Rodríguez Mateos, D. (2007). Herramientas de software para OAI-PMH.

Recuperado de:

<http://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0C CkQFjAC&url=http%3A%2F%2Fwww.uv.es%2F%3Dbarrueco%2Fcardedeu.doc&ei=2kU0VMryGpSlyATZ8YKABg&usg=AFQjCNFtMxHwtgEgjFD6D-iLstdwZ5g63A&bvm=bv.76943099,d.b2U>

Bueno de la Fuente y Rodríguez Mateos, D. (2007). La búsqueda de alternativas de acceso a los documentos científicos.

Recuperado de:

<http://e-archivo.uc3m.es/bitstream/10016/9084/5/CAPITULO1-OAI-troduccion.pdf>

Bustos González, A., Fernández Porcel A. y Johnson, I. (2007). Directrices para la creación de repositorios institucionales en universidades y organizaciones de educación superior.

Recuperado de:

http://eprints.rclis.org/16746/1/Directrices_RI_Espa_ol.pdf

Candás Romero, J. (2006). El papel de los metadatos en la preservación de la información

Recuperado de:

<http://eprints.rclis.org/8359/1/final.pdf>

Chan L. (2004). Supporting and enhancing scholarship in the digital age : the role of Open Access institutional repositories. *Canadian Journal of Communication Corporation*. 29, 277-300.

Chapman, J. W., Reynolds, D. y Shreeves, S. A. (2009). Repository metadata : approaches and challenges. *Cataloging & Clarification Quaterly*, 47, 309-325.

Civallero, E. (2008). Open Access : experiencias latinoamericanas. Recuperado de:
http://eprints.rclis.org/bitstream/10760/11621/1/Open-Access-Experiencias_latinoamericanas.pdf

Codina, L. (2009). Sistemas de información y metadatos en la web semántica. En: *Web semántica y sistemas de información documental*. (p. 13-42). Madrid : Ediciones Trea.

Comunicación formal e informal (2013).
Recuperado de:
http://www.erevistas.csic.es/especial_revistas/revistas1.htm

Conyers, J. (2009). John Conyers and Open Access.
Rcuperado de:
<http://www.lessig.org/blog/2009/o3/john-conyers-and-open-access/html>

Córdoba González, S. (2010). Los repositorios de acceso abierto : una vía para la democratización de la información y el conocimiento. *Seminario de Bibliotecología, información y sociedad. Derecho a la información, bien público y bien privado: acceso comunitario y acceso individual*, 29-56.

Caribé, Rita de Cassia do Vale (2008). A aplicação do desenvolvimento e gerenciamento de coleções na construção de repositórios institucionais.
Recuperado de:
<http://www.brapci.ufpr.br/documento.php?dd0=0000005038&dd1=806ca>

Crow, R. (2003). The case of Institutional Repositories : a SPARC position per.
Recuperado de: http://www.sparc.arl.org/sites/default/files/media_files/instrepo.pdf

La definición de software libre (2013).
Recuperado de:
[http://ecaths1.s3.amazonaws.com/nticx7/883846354.La%20Definici%C3%B3n%20de%20Software%20Libre%20-%20Proyecto%20GNU%20-%20Free%20Software%20Foundation%20\(FSF\).pdf](http://ecaths1.s3.amazonaws.com/nticx7/883846354.La%20Definici%C3%B3n%20de%20Software%20Libre%20-%20Proyecto%20GNU%20-%20Free%20Software%20Foundation%20(FSF).pdf)

Dempsey, L. y Heery, R. (1998). Metadata : a current view of practice and issues. *Journal of documentation*. Marzo 54 (2), 145-172.

Estivill, A., Abadal, E., Franganillo, J., Gascón, J. y Rodríguez Gairín, J. M. Uso de metadatos Dublin Core en la descripción y recuperación de artículos de revistas digitales = use of Dublin Core metadata describing digital journals.
Recuperado de:

<http://eprints.rclis.org/6760/1/estivill2005es.pdf>
Flores Cuesta, G. (2007). Los repositorios institucionales : análisis de la situación internacional y principios generales para Cuba.
Recuperado de:
http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352007001200006

García, N.E., Caballero, S. (2009). Metadatos : necesidad e importancia de crear estándares
Recuperado de:
http://www.bn.gov.ar/descargas/catalogadores/ponencias/251109_10a.pdf

Goodchild, M. F. (1995). Alexandria digital library : report workshop on metadata.
Recuperado de:
http://www.legacy.alexandria.ucsb.edu/public-ocuments/metadata/metadata_ws.html

Goodman, D. (2004). The criteria for Open Access.
Recuperado de:
http://eprints.rclis.org/6167/2/SR30_4_Goodman.pdf

Hanrad, S. (2001). Six proposals for freeing the refereed literature online : a comparison. Recuperado de:
<http://cogprints.org/1702/>

Hanrad S. (2012). What is Open Access?.
Recuperado de:
<http://www.eprints.org/openaccess/>

Hanrad, S., Les Carr, B., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., Hajjem, C. y Hilf, E. (2008). The Access/impact problem and the green and gold roads to Open Access. *Serials review*. 34, 36-40.

Haynes, D. (2004). *Metadata for information management and retrieval*. London : Faced publishing.

Hixson, C. y Cracknell, Linda (2007). How to implement an institutional repository.
Recuperado de:
http://www.tandfonline.com/doi/abs/10.1300/J123v52n01_05#preview

Hunter, P. y Day, M. (2005). Institutional repositories, aggregator services and collection development.
Recuperado de:
<http://eprints-uk.rdn.ac.uk/project/docs/studies/coll-development/coll-development.pdf>

La iniciativa de Acceso (2002.). Budapest Open Access Initiative (BOAI).
Recuperado de:

<http://www.madrimasd.org/informacionidi/e-ciencia/documentos/docs/BOAI.pdf>

Keefer, A. (2003). Aproximació al moviment "open Access".

Recuperado de:

<http://bid.ub.edu/15keefer.htm>

Ley Federal del Derecho de autor

Recuperado de

http://www.indautor.gob.mx/documentos_normas/leyfederal.pdf

Lima Leite, F. C. y Costa, S. (2006). Repositórios institucionais como ferramentas de gestão do conhecimento científico no ambiente acadêmico.

Recuperado de:

<http://www.scielo.br/pdf/pci/v11n2/v11n2a05.pdf>

Lynch, Clifford, A. (2011). Creación de un repositorio digital con recursos limitados. Recuperado de:

<http://docudigital.wordpress.com/2011/06/15/creacion-de-un-respositorio-digital-con-recursos-limitados/>

Lynch Clifford, A. (2003). Institutional repositories: esencial infrastructure for scholarship in the digital age.

Recuperado de:

<http://www.arl.org/storage/documents/publications/arl-br-226.pdf>

Martínez Arellano, F. F. (2012, octubre). El acceso abierto en América Latina. *Ponencia presentada en el marco de la Semana Internacional de Acceso Abierto (Open Access Week)*. UNAM, IIBI, México, D.F.

Recuperado de:

http://www.youtube.com/results?search_query=conferencia+Mart%C3%ADnez+arellano

Martínez Arellano, F. F. (coord.)(2011). *Coloquio acceso abierto a la información en las bibliotecas académicas de América Latina y el Caribe*. México, UNAM, CUIB, 348 p.

Martínez Arellano, F. F. (2011). Presentación. En: *Coloquio acceso abierto a la información en las bibliotecas académicas de América Latina y el Caribe*. México : UNAM, CUIB, p. ix-xii.

Melero, R. (2007). Acceso Abierto a las publicaciones científicas : definición, recursos, copyright e impacto.

Recuperado de:

<http://eprints.rdis.org/6571/1/EPI-rmelero.pdf>

Melero, R. (2006). Open Access y repositorios institucionales.

Recuperado de:

http://www.iata.csic.es/~bibrem/OPEN_ACCESS/repositorios-articulomelero.pdf

Melero, R. (2004). Plataforma digital de revistas científicas electrónicas españolas. Relación con el movimiento Open Access.

Recuperado de:

<http://digital.csic.es/bitstream/10261/1/OA4m.pdf>

Melero, R. y Abad García, M. F. (2008). Revistas open Access : características, modelos económicos y tendencias.

Recuperado de:

<http://bid.ub.edu/20meler2.htm>

Méndez Rodríguez, M. (2002). *Metadatos y recuperación de información : estándares, problemas y aplicabilidad en bibliotecas digitales*. Madrid : Ediciones Trea.

Morales Campos, E. y Martínez Arellano, F. F. (2011). Acceso Abierto a la producción científica en humanidades y ciencias sociales en la Universidad Nacional Autónoma de México.

Recuperado de:

<http://conference.ifla.org/past-wlic/2011/104-campos-es.pdf>

Moyano Grimaldo, W. A. (2004). Sociedad de la información : metadatos y futuro de la internet en la recuperación de información de calidad.

Recuperado de:

http://eprints.rdis.org/7010/1/Art%C3%ADculo_Sociedad_de_la_Informaci%C3%B3n_y_Metadatos.pdf

Mu, E. (2009). Open Access electronic journals : issues and oportunities for Latin America. Presentación para el *Congress of the Latin American Studies Associations*. Brasil, julio 11-14 Rio de Janeiro.

Odlyzco, A. (2009). Chapter 4 : economic costs of toll Access.

Recuperado de:

http://www.dtc.umn.edu/~odlyzko/doc/OpenAccess_book-odlyzko-chapter.pdf

Open Archives Initiative : frequently asked questions (FAQ) (2002).

Recuperado de:

<http://www.openarchives.org/documents/FAQ.html>

OpenDOAR. directory of open access repositories (2014).

Recuperado de:

<http://www.openoar.org/find.php>

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (2010). *Greenstone : un software libre de código abierto para la construcción de bibliotecas digitales : experiencias en América Latina y el Caribe*. Montevideo : UNESCO.

Recuperado de:

<http://unesdoc.unesco.org/images/0018/001887/188719s.pdf>

Passini Moreno, F., Lima Leite, F.C. y Márdero Arellano, M. (2006). Acceso libre a publicaciones e repositorios digitais em ciencia da informação no Brasil.

Recuperado de:

<http://www.scielo.br/pdf/pci/v11n1/v11n1a07.pdf>

Repository applications (2008).

Recuperado de:

<http://questia.com/library/journal/101-175020813/repository-applications>

Public record office the National Archives (2002). Requirements for electronic records management systems. 2 : metadata standard.

Recuperado de:

<http://www.nationalarchives.gov.uk/documents/metadatabfinal.pdf>

Richard, J. Roberts, et al. (2001). Building a “gen bank” of the published literature. *Science* 291 (5512), 2318-2319.

Rodríguez Bravo, B. (2007). Los repositorios de información, guardianes de la memoria digital. *Anales de documentación*. (10), 361-374.

Rodríguez Gallardo, A. (2008). Elementos que fundamentan el Acceso Abierto. *Investigación bibliotecológica*. 22 (44), 161-182.

Rucinque, H. F. (2003). Más sobre acceso abierto.

Recuperado de:

http://www.madrimasd.org/informacionidi/e-ciencia/documentos/docs/declaration_berlin.pdf

Sánchez García de las Bayonas, S. y Melero, R. (2007). La denominación y el contenido de los repositorios institucionales en acceso abierto : base teórica para la ruta verde.

Recuperado de:

http://eprints.rdis.org/bitstream/10760/10367/1/PFC_Merce_Millet.pdf

Sánchez Tarragó, N. (2009). El movimiento de Acceso Abierto a la información y las políticas nacionales e institucionales de autoarchivo.

Recuperado de:

http://bvs.sld.cu/revistas/aci/vol16_3_07/aci05907.html

Santamarina de la Vega, Carolina (2013). Metadatos y otros temores.
Recuperado de:
http://www.academia.edu/6574246/Metadatos_y_otros_temores

Sanz-Valero, J., et al. (2007). La iniciativa Open Access, una visión de conjunto.
Recuperado de:
http://scielo.isciii.es/scielo.php?pid=S0465-546X2007000200003&script=sci_arttext

Sarduy Domínguez, Y. y Urra González, P. (2006). Herramientas para la creación de colecciones digitales.
Recuperado de:
http://scielo.sld.cu/scielo.php?pid=S1024-94352006000500019&script=sci_arttext

Science News (2012).
Recuperado de:
(<http://www.sciencedaily.com/releases/2012/07/120716214919.htm>).

Senso, J. A. y de la Rosa Piñero, A. (2003). El concepto de metadato : algo más que descripción de recursos electrónicos. *Ciência da informação*, 32 (2), 95-106.

Serrano, J. y Prats, J. (2005). Repertorios abiertos : el libre acceso a los contenidos. Recuperado de:
<http://www.uoc.edu/rusc/2/2/dt/esp/serrano.pdf>

Spano, Diego J. (2010). El Open Source como facilitador del Open Access.
Recuperado de:
<http://dialnet.unirioja.es/servlet/articulo?codigo=3342024>

Suber, P. (2006). Una introducción al acceso abierto.
Recuperado de:
https://www.zotero.org/groups/acceso_abierto/items/itemKey/WP2QQJ3K

Suber, P. and Subbiah Arunachalam (2006). Open Access to Science in the Developing World. in *Open Access: Unrestricted Access to Published Research*. Sahu, D. K. ed. 93rd Indian Science Congress, 3-7 Jan. 2006, Hyderabad, India.
http://openmed.nic.in/1359/01/OA_ISC.pdf#page=4

Taylor, A. G. (2004). *The organization of information*. London : Libraries.

Toricela Morales, R. G., Lee Tenorio, F. y Huerta Espinosa, V. M. (2008). Acceso abierto y software libre : premisas para la independencia tecnológica = Open Access and free software : premises for technological Independence.
Recuperado de:

http://bvs.sld.cu/revistas/aci/vol17_2_08/aci09208.htm

Tramullas, J. (2014). Software libre para exposiciones digitales.

Recuperado de:

<http://tramullas.com/software-libre-para-exposiciones-digitales/>

Tramullas, J., Garrido Picazo, P. (2006). Software libre para repositorios institucionales : propuestas para un modelo de evaluación de prestaciones. *El profesional de la información*, 15 (3), 171-181.

University of Southampton (2012). EPrints.

Recuperado de:

<http://www.eprints.org/>

Vázquez Paulus, C. 2002? Metadatos : introducción e historia.

Recuperado de:

<http://users.dcc.uchile.cl/~cvasquez/introehistoria.pdf>

Vinicio Caballero, C., et al. (2008). La importancia del acceso abierto en la investigación biomédica y científica.

Recuperado de:

http://www.scielo.org.co/scielo.php?pid=S0121-81232008000200004&script=sci_arttext

Witten, I. H., et al. (2005). A bridge between Greenstone and DSpace. *D-Lib Magazine*.

Recuperado de:

<http://www.dlib.org/dlib/september05/witten/09witten.html>