

Directrices para la construcción de un repositorio temático

Maria Vanesa Doria¹, Claudia Inés Inchaurredo¹, Germán Montejano^{2,3}

¹ Dpto. Sistemas. Facultad de Tecnología y Cs. Aplicadas. Universidad Nacional de Catamarca. Maximio Victoria 55. CP 4700. San Fernando del Valle de Catamarca. Argentina.

² Facultad de Cs. Físico, Matemáticas y Naturales, Universidad Nacional de San Luis, Ejército de los Andes 950 – CP: 5700 – San Luis

³ Facultad de Ingeniería, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa – La Pampa

Resumen

La creación de repositorios de información institucionales y/o temáticos, en nuestros días es la clave para el acceso a producción científica de manera libre y abierta.

Por lo tanto, el acceso a la información a través de repositorios de información se ve favorecido por las iniciativas que propone el Acceso Abierto (AA).

Para la creación de un repositorio de información en el marco de la filosofía de AA es importante contar con directrices, éstas pretenden ser una guía general de aspectos básicos donde se detallan los pasos necesarios para la creación de un repositorio.

El presente trabajo recoge las directrices para la construcción de repositorios temáticos a partir de los siguientes puntos de referencia: finalidad, funciones, políticas y actividades.

Aunque existe un consenso sobre que los repositorios temáticos fueron los primeros en aparecer, existe poca información sobre desarrollo y gestión de repositorios temáticos comparado con la amplitud de información alcanzada por repositorios institucionales.

Palabras clave: repositorios temáticos, directrices, acceso abierto, finalidad, funciones, políticas, actividades.

Abstract

The creation of institutional or thematic information repositories, in our day is the key to access to scientific production freely and openly.

Therefore, access to information through information repositories is favored by the initiatives proposed by the Open Access (AA).

To create a repository of information in the context of the philosophy of AA is important to have guidelines, they are intended as a general guide to the basic aspect which details the steps necessary to create a repository.

This paper presents guidelines for building thematic repositories from the following reference points: purpose, functions, policies and activities.

Although there is consensus about that the thematic repositories were the first to appear, there is little information about the development and management of thematic repositories compared with the breadth of information achieved by institutional repositories.

Keywords: thematic repositories, guidelines, open access, purpose, functions, policies, activities.

1. Introducción

El Diccionario de la Real Academia (DRAE) define un repositorio como “Lugar donde se guarda algo”, por otra parte, el término deriva del latín “repositorium”, que significa armario o alacena, partiendo de estas definiciones se aplicó al léxico específico de la informática para designar los repositorios de información digital.

Un repositorio de información digital es un sistema de red formado por hardware, software, datos y procedimientos. Este sistema sirve para almacenar, conservar y dar acceso a documentos digitales.[1]

Se pueden nombrar las siguientes características en lo que se refiere a repositorios de información, las cuales han sido tomadas de distintos autores, tales como, [2][3][4]:

- Grandes volúmenes de datos que contienen documentos de distintos tipos.
- Objetos digitales fácilmente accesibles a través de una estructura estandarizada de metadatos y

recuperables para ser usados, reutilizados y compartidos con diferentes propósitos y en diferentes contextos.

- Operados por universidades u organizaciones de investigación.
- La unidad básica y única para obtener información es el documento.
- Preparados para distribuirse habitualmente sirviéndose de una red informática (ejemplo: Internet ó Intranet) o en un medio físico (ejemplo: CD).

El acceso a la información a través de repositorios de información se ve favorecido por las iniciativas que propone el Acceso Abierto (AA).

El AA es una filosofía en la que los usuarios pueden leer, descargar, copiar, distribuir, imprimir, buscar, o enlazar los textos completos de los artículos científicos, y, usarlos con cualquier otro propósito legítimo, sin otras barreras financieras, legales o técnicas más que las que suponga Internet en sí misma. La única restricción que se impone sobre la distribución y reproducción es la de dar a los autores control sobre la integridad de su trabajo y el derecho a ser citado y reconocido adecuadamente [5].

Por lo tanto un repositorio de AA es un repositorio de información digital que recoge, almacena y da acceso a resultados (publicaciones y/o documentos) de investigación científica.[1]

Un repositorio de AA representa una alternativa a los métodos tradicionales de publicación y difusión de la producción científica; si bien no es un canal formal de publicaciones sirve como estrategia para acceder a la información de manera libre y abierta. Libre se refiere a la consulta del documento completo en línea sin costo y abierta esta relacionada con algunos derechos de uso adicionales como el de modificar y distribuir la obra siempre y cuando se cite a él o los autores

Habitualmente, los repositorios de AA se clasifican en repositorios institucionales y repositorios temáticos. Se habla de repositorios institucionales cuando los objetos digitales que contienen reflejan la producción científica de una organización (universidad, departamento, instituto, sociedades científicas); los temáticos, están referidos al depósito de objetos digitales vinculado a una disciplina, según Lynch [6], abrieron camino en el ciclo de comunicación científica, esto permitió a las comunidades académicas identificar el potencial de esta alternativa de acceso a la literatura científica. [7]

Si bien los objetivos específicos de un repositorio institucional con respecto a un repositorio temático pueden variar, en líneas generales podrán resumirse en los siguientes puntos:

- Recoger y preservar los resultados de la investigación en formato digital, garantizando su permanencia en el tiempo.
- Abrir y ofrecer a la comunidad científica y a la sociedad en general los resultados de las investigaciones realizadas por miembros de la

institución o de aquellos que trabajan en un determinado ámbito temático.

- Aumentar el impacto y la influencia de los resultados, maximizando la visibilidad de los resultados.
- Ser un repositorio interoperable ya que de este modo, permite el intercambio de documentación científica y el acceso a la misma de una manera rápida y potencialmente sencilla.

Por otra parte, las fronteras entre repositorios y bibliotecas digitales u otros servicios similares resulta difusa, por que en la literatura existente estos términos se han utilizado como sinónimos, prefiriendo uno u otro en función del ámbito en que se desarrolle la investigación.

En el caso de producciones científicas de una disciplina en particular predomina la denominación “repositorios temáticos”, no existe unanimidad conceptual sobre esta denominación, hay aportes de diferentes de autores que coinciden que es una colección de objetos digitales (documentos y publicaciones) cuyo eslabón principal común es el campo disciplinar. Las características que se pueden mencionar son [8]:

- Permite el acceso al texto completo.
- El depósito es voluntario y se conoce como proceso de autoarchivo.
- La comunidad de pares asume la gestión del contenido.
- La práctica de la moderación es un alternativa, pero no es una práctica común.
- La meta y el compromiso es la divulgación del conocimiento científico.

En el siguiente apartado se describen consideraciones generales sobre los repositorios temáticos.

2. Repositorios Temáticos

Existe un consenso general respecto a que los repositorios temáticos fueron los primeros repositorios en aparecer, pueden describirse como proveedores de servicios que recolectan datos estructurados relacionados con un tema, permitiendo posteriormente al usuario buscar a través de muchas fuentes distribuidas en todo el mundo desde un punto de acceso único [4][5].

Se desarrollaron en el ámbito de disciplinas académicas concretas, tales como, la física, ciencias de la información, ciencias cognitivas, salud, entre otras.

2.1. Características Principales

Las características principales de los repositorios temáticos, son las siguientes [1]:

- Permiten almacenar documentos pre-prints (antes de pasar cualquier proceso de arbitraje) o post-prints (documentos revisados por pares y aceptados,

publicados o en proceso de publicación) y pueden ser artículos de revistas, comunicaciones en congresos, capítulos para libros o cualquier otra forma de comunicación científica.

- Los documentos disponibles en los repositorios de información se encuentran a texto completo y el acceso es gratuito, libre de cualquier tipo de restricciones.
- Los propios autores son los responsables de la introducción de los documentos en el repositorio, mediante el proceso de autoarchivo.

2.2. Tratamiento de Documentos

Autoarchivo

El término autoarchivo se refiere al proceso que le permite al autor de un documento introducirlo en el repositorio en el que se encuentre registrado.

En el espacio destinado para subir los documentos introducirá los metadatos, descripciones obligatorias y opcionales que están disponibles de forma pública para ayudar a localizar objetos, en las áreas de telecomunicaciones e informática, es la información no relevante para el usuario final pero de suma importancia para el sistema que administra los documentos.

Metadatos

Etimológicamente la palabra metadatos significa “algo más que datos”, es decir, constituyen información sobre los datos [9].

Los metadatos de un recurso de información que se encuentra en un repositorio son datos descriptivos, en el caso de trabajos científicos son título, palabras clave, resumen, fecha, entre otros que permiten la identificación y su localización a través de la computadora. Estos datos son recolectados por diferentes proveedores de datos (administradores de contenido) que soportan el protocolo Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).

OAI-PMH, es un protocolo para la definición e intercambio de metadatos de un archivo, esto significa que los archivos que cumplen con las iniciativas del AA se inscriben en un registro central y pueden ser tratados como un único y gran archivo virtual con propósitos de búsqueda. Este protocolo surge en la Convención de Santa Fé, (Nuevo México, Estados Unidos) en julio de 1999. [10]

Las características de este protocolo son [11]:

- Simplicidad: los creadores buscaron una forma simple que se encuentre al alcance de cualquier usuario.
- Normalización: basados en estándares ampliamente utilizados como el protocolo http (Hiper Text Transport Protocol) para la transmisión de datos y

órdenes y XML (Extended Markup Language) para la codificación de metadatos.

- **Recolección:** existe una entidad que pone a disposición de los interesados información bibliográfica sobre los documentos que almacena.

Por otro lado, los elementos del protocolo son los siguientes [12]:

- **RECURSO:** Es el documento o producción científica
- **ITEM:** Colección de propiedades sobre un RECURSO, es decir, metadatos que describen el recurso (título, autor, etc.)
- **REGISTRO:** Es un ITEM expresado en un formato específico formado por:
 - **Encabezamiento:** contiene un identificador único (identifica de forma inequívoca un artículo dentro de un repositorio), marca de fecha (fecha de creación, modificación o eliminación del registro con el fin de la recolección selectiva) cero o el conjunto de miembros de la partida con el propósito de la recolección selectiva, un opcional de estado de un atributo con un valor de borrado que indica el retiro de la disponibilidad del formato de metadatos especificado para el elemento.
 - **Metadatos:** se debe especificar le modelo de metadatos elegido
 - **Acerca de:** es una parte opcional en la que se pueden definir declaraciones de derechos y declaraciones de procedencias.

El protocolo recomienda el modelo Dublín Core sin cualificar que se compone de 15 elementos agrupados en las siguientes funciones: contenido, propiedad intelectual, instanciación. [13]

3. Aspecto Legal

Un tema que reviste importancia en el ámbito de los repositorios es el conocimiento de los aspectos legales referidos a la propiedad intelectual, los derechos de autor y las licencias.

Sobre propiedad intelectual se debe conocer lo establecido en Argentina por la ley de Propiedad Intelectual y el concepto de la Organización Mundial de la Propiedad Intelectual (OMPI).

- Según la Ley de Propiedad Intelectual 11.723 [14] del Poder Ejecutivo de la Nación Argentina, “el autor de obras científicas, literarias y/o artísticas tiene la facultad de disponer de ella, de publicarla, de ejecutarla, de representarla, y exponerla en público, de enajenarla, de traducirla, de adaptarla o de autorizar su traducción y de reproducirla en cualquier forma.”
- El concepto sobre propiedad intelectual que brinda la Organización Mundial de la Propiedad Intelectual (OMPI) es [15]: “La propiedad intelectual se refiere a las creaciones de la mente:

invenciones, obras literarias y artísticas, así como símbolos, nombres e imágenes utilizadas en el comercio. La propiedad intelectual se divide en dos categorías:

- La propiedad industrial, que incluye las patentes de invenciones, las marcas, los diseños industriales y las indicaciones geográficas.
- El derecho de autor, que incluye obras literarias, tales como novelas, poemas y obras de teatro, películas, obras musicales, obras artísticas, tales como dibujos, pinturas, fotografías y esculturas, y diseños arquitectónicos. Los derechos conexos al derecho de autor incluyen los derechos de los artistas intérpretes o ejecutantes sobre sus interpretaciones o ejecuciones, los de los productores de fonogramas y los de los organismos de radiodifusión respecto de sus programas de radio y televisión.”

En lo referente a derecho de autor destacamos el marco que existe en nuestro país y lo expresado por OMPI a nivel internacional:

- En Argentina, el Derecho de autor está enmarcado, en principio, por el artículo 17 de la Constitución que expresa que: “Todo autor o inventor es propietario exclusivo de su obra, invento o descubrimiento, por el término que le acuerde la ley. La Ley 11.723 [14] regula el régimen Legal de la Propiedad Intelectual. El artículo 5 de la Ley de Propiedad Intelectual (Argentina) dice que “La propiedad intelectual sobre sus obras corresponde a los autores durante su vida y a sus herederos o derechohabientes hasta setenta años contados a partir del 1 de enero del año siguiente al de la muerte del autor.” En el caso de obras realizadas en colaboración, el plazo se cuenta desde el 1 de enero del año siguiente a la muerte del último de los autores. Si el autor no dejara herederos, los derechos pasan directamente al Estado Argentino por el mismo plazo que estipula la ley.
- La OMPI [15] expresa: “El derecho de autor es el cuerpo de leyes que concede a los autores, artistas y demás creadores protección por sus creaciones literarias y artísticas, a las que generalmente se hace referencia como “obras”. El titular o titulares de una obra pueden prohibir u autorizar:
 - su reproducción bajo distintas formas, incluida la publicación impresa y la grabación sonora;
 - su interpretación o ejecución pública, así como su comunicación al público;
 - su radiodifusión;
 - su traducción a otros idiomas; y
 - su adaptación.”

Relacionado con el derecho de autor se encuentran las licencias que determinan las modalidades de utilización de una obra.

Una licencia es un instrumento legal que indica qué se puede y qué no se puede hacer con una obra intelectual. Toda licencia presenta tres partes:

- Condiciones a cumplir.
- Derechos cedidos por el titular del Copyright.
- Limitación de responsabilidad.

Feltrero Oreja [16] propone el uso de licencias libres para la protección de los usos legítimos de la producción intelectual, “especialmente si se trata de conocimiento científico o humanístico subvencionado públicamente”. “Copyleft” es el término usado para referirse a las licencias libres.

La Fundación Copyleft [17] define Copyleft como “grupo de licencias cuyo objetivo es garantizar que cada persona que recibe una copia de una obra pueda a su vez usar, modificar y redistribuir el propio trabajo y las versiones derivadas del mismo. Unas veces se permite el uso comercial de dichos trabajos y en otras ocasiones no, dependiendo qué derechos quiera ceder el autor.”

Como licencias Copyleft podemos mencionar [18]:

- Licencias Creative Commons o CC (en español: «bienes comunes creativos») están inspiradas en la licencia GPL (General Public License) de la Free Software Foundation. Pertenecen a la organización no gubernamental sin fines de lucro de Creative Commons.
- Licencias de Documentación Libre de GNU (FLD GNU), estas licencias son una forma de copyleft para ser usada en un manual, libro de texto u otro documento que asegure que todo el mundo tiene la libertad de copiarlo y redistribuirlo, con o sin modificaciones, de modo comercial o no comercial.

4. Directrices para la creación de un repositorio temático

Las directrices que se describen en esta sección pretenden ser una guía general de aspectos básicos donde se detallan los pasos para la creación de un repositorio temático, abordándose para ello los siguientes puntos de referencia: finalidad, funciones, políticas y actividades.

4.1. Finalidades

Las finalidades de un repositorio temático son compartir, reutilizar y preservar resultados de investigación en un campo disciplinar.

A continuación se realiza una breve descripción de cada finalidad:

- Compartir está referido a fomentar la colaboración y el intercambio de producciones en el campo disciplinar, se debe decidir que tipo de producciones se tendrán en cuenta: experiencias, artículos, ponencias que se hayan presentado en algún evento o revista científica, entre otras.

- Reutilizar contribuye a evitar el tener que “reinventar la rueda” continuamente, permite generar nuevos contextos de investigación, así como servir de base para la creación de nuevas producciones.
- La preservación digital persigue dos objetivos [19]:
 1. que los objetos puedan ser mantenidos en el repositorio sin ser dañados, perdidos o alterados maliciosamente;
 2. que los datos puedan ser encontrados, extraídos de los archivos y ofrecidos a los usuarios, para que puedan ser interpretados y entendidos por estos.

4.2. Funciones

El repositorio debe ofrecer un conjunto de funciones de utilidad para los administradores, gestores y usuarios, estas funciones se pueden organizar en dos niveles:

- Nivel técnico y de contenidos: están relacionadas con gestionar el contenido, los usuarios y los procesos internos del sistema, y donde principalmente serán necesarios interfaces de administrador técnico y de gestor de contenidos, pero donde también tienen lugar los usuarios contribuyentes de contenido que accederán a algunas de las funcionalidades de depósito.
- Nivel de servicios y recursos: están relacionadas con buscar, seleccionar y obtener documentos tanto si son usuarios registrados como si no; serán necesarios múltiples interfaces para los distintos roles que se hayan definido.

Nivel técnico y de contenidos

Las funciones en este nivel se refieren a:

- Gestión técnica: se encuentran funciones sobre presentación de datos e información (formato de archivos, extensiones de archivos, formatos de comprensión, codificación de fecha y hora)
- Gestión de arquitecturas de redes y comunicación: asociada a protocolos de Internet, seguridad y servicios.
- Gestión de almacenamiento: consiste en definir formatos soportados y preferidos, asignación de identificadores únicos, exportación e importación de documentos.
- Gestión de usuarios: referidos a altas, bajas, modificaciones y perfiles de usuarios para definir distintos niveles de edición, administración y acceso.
- Control de acceso: relacionado con definir niveles y derechos de acceso distintos según los grupos y perfiles de usuarios, mecanismos de autenticación y registro de sesiones de usuarios.
- Gestión de contenidos: consiste en definir depósito de archivos, gestión de metadatos, gestión de derechos y preservación.

Nivel de servicios y recursos

Las funciones en este nivel se refieren a:

- Búsqueda: ofrecer múltiples opciones de búsqueda efectivas y de calidad, así como índices de navegación que ofrezcan múltiples puntos de acceso a los documentos, ofrecer opciones de búsqueda simple y avanzada, en la que se permita buscar tanto por texto libre como por distintos elementos de metadatos de los registros, búsqueda a texto completo, empleo de operadores booleanos, truncamiento o proximidad, así como diversas opciones de presentación, ordenación y clasificación de resultados. Los elementos de metadatos básicos para la búsqueda incluyen aspectos de identificación (título, autor), contenido (descripción, palabras clave, clasificación), es decir, ofrecer opciones de búsqueda simple y avanzada. La potencialidad del motor de búsqueda que se emplee en el repositorio determinará la capacidad y facilidad de éste para recuperar su contenido.
- Acceso y previsualización: permitir el acceso tanto al registro de metadatos del documento como al propio documento. El registro de metadatos debe mostrar al usuario todo tipo de información de relevancia que le pueda ayudar en la decisión de seleccionar o no un objeto educativo. Una funcionalidad interesante a incorporar en un repositorio es la previsualización como visores de PDF.
- Formatos de descargas: ofrecer múltiples opciones de descarga para que los usuarios seleccionen la que más se adecúe a sus necesidades o preferencias (html, pdf, Word, rtf, etc.)

4.3. Políticas

Resulta interesante tener en cuenta lo expresado por Bueno de la Fuente [20]: Un marco general de políticas es una herramienta vital a través de la cual establecer los límites operacionales en los que funcionará el repositorio, planificando el día a día con una visión de futuro a largo plazo. Los beneficios fundamentales de establecer un marco de políticas claro y visible son: facilitar el entendimiento de los usuarios y otras partes interesadas; ayudar a la planificación y a los procesos de tomas de decisiones; y asegurar el soporte del repositorio.

Las políticas que se recomiendan definir son: políticas de contenido, políticas de metadatos, políticas de depósito, políticas de preservación de documentos y políticas de propiedad intelectual y licencias

Políticas de Contenido

Para las políticas de contenidos se deberá especificar si se almacenan los documentos que vayan a quedar en acceso libre y/o los documentos que hayan sido o vayan

a ser publicados por canales formales, para las tipologías de contenidos las más conocidas son las siguientes:

- Tesis doctorales terminadas.
- Comunicaciones a congresos (ponencias, posters, etc).
- Pre-prints y post-prints.
- Software.

Se debe decidir también si el repositorio gestionará sólo contenidos almacenados físicamente en el repositorio o si se aceptarán referencias a objetos externos.

Sobre el formato de los documentos se debe tener en cuenta el sistema de software que se utilizará, se puede crear una lista de formatos organizados por categorías:

- Aceptados: soporte completo al formato.
- Conocidos: se reconocen pero no se garantiza el soporte completo.
- No aceptados: el sistema no reconoce el formato.

Sobre el tamaño de los documentos se puede optar por no aceptar documentos de gran tamaño, en gran parte la restricción de tamaño está relacionada con el sistema de software elegido.

Políticas de Metadatos

Las principales decisiones que se deben tomar son:

- Modelo de metadatos que se utilizará, o si se desarrollará un modelo de metadatos personalizado.
- Determinar si habrá metadatos obligatorios.
- Determinar si el elemento que rellenará el metadato debe estar de acuerdo a alguna norma, por ejemplo 610 - IEEE Standard Glossary of Software Engineering Terminology,
- Establecer si para los términos más específicos se utilizará un tesoro para aportar contenido temático y evitar así el uso de variantes.

Políticas de Depósito

La política de depósito debe distinguir si se permite la participación de personas externas o solo de los miembros del repositorio, también debe establecer las formas de depósito, que pueden ser [21]:

- Depósito directo: Persona X una vez ingresado al repositorio como usuario autorizado carga el archivo y lo publica en el repositorio.
- Depósito mediado por un administrador: Persona X envía por correo electrónico el archivo y la información, Persona Y (Administrador) confirma origen, una vez ingresado al repositorio como usuario autorizado carga el archivo y lo publica en el repositorio.
- Depósito semi-mediado: Persona X una vez ingresado al repositorio como usuario autorizado carga el archivo, lo remite a una carpeta, Persona Z (Revisor) revisa los depósitos pendientes, confirma, mejora, completa los datos necesarios, acepta o rechaza la publicación.

Políticas de Preservación de documentos

Estas políticas están relacionadas con las decisiones de:

- Determinar qué se preserva y durante qué periodo (medio o largo plazo).
- Definir qué aspectos de los documentos se preservarán: la funcionalidad original, la apariencia y el diseño original, la secuencia de bits o ficheros, los metadatos, las anotaciones de usuarios, etc.
- Se debe definir también la forma de realización de copias de seguridad, en qué ubicación (servidor externo o no) y con qué frecuencia.

Políticas de Propiedad intelectual y licencias

Esta política debe adaptarse a la legislación nacional e internacional vigente.

La política del repositorio debe contemplar al menos dos tipos de licencias:

1. De autor, la que acepta el autor al depositar sus recursos en el repositorio.
2. De uso, la que acepta el usuario al acceder a los recursos.

Además, es recomendable contar con una política o declaración general del repositorio que el usuario se compromete a respetar por el mero hecho de acceder al repositorio y utilizar sus recursos.

Políticas de usuarios

El repositorio debe distinguir por lo menos estos grupos:

- Usuarios Finales: tendrán un nivel de acceso público que les permitirá ver y descargar los documentos sin ser un usuario registrado.
- Usuarios Registrados: que puede tener el rol de autor de documentos, administrador del repositorio o revisor de documentos y tienen un nivel de acceso privado, puede limitarse a espacios personales o en grupo de determinados materiales que no se desee poner a disposición del resto de usuarios por distintas razones.

4.4. Actividades para la creación de un repositorio temático

A modo orientativo se ofrece el siguiente flujo de actividades básico para la construcción de un repositorio temático de acuerdo a las políticas detalladas anteriormente:

1. Análisis documental del cuerpo de documentos que formaran parte del repositorio.
2. Análisis de requerimientos.
3. Diseño.
4. Selección de tecnología e implementación.

Análisis documental del cuerpo de documentos que formaran parte del repositorio

Primero se debe elegir la disciplina temática y las diferentes subdisciplinas que determinarán el alcance temático del repositorio.

El análisis del cuerpo documental del repositorio supone realizar las operaciones de identificación, descripción formal conforme al modelo metadatos elegido, organización y clasificación en el repositorio.

Para la operación de clasificación se puede utilizar instrumentos vigentes aceptados internacionalmente para establecer el alcance temático, como por ejemplo tesauros.

Los tesauros son una lista controlada y estructurada de términos para el análisis temático y ayudan a la búsqueda de documentos y publicaciones. La Organización de las Naciones Unidas para la Educación, Ciencia y la Cultura (UNESCO) ofrece un tesauro para los campos de la educación, cultura, ciencias naturales, ciencias sociales y humanas, comunicación e información, continuamente ampliada y actualizada. [22]

Análisis de requerimientos

Un requerimiento es una condición o funcionalidad que debe ser satisfecha o poseída por un sistema o por alguno de sus componentes [23].

Para el caso de la construcción de un repositorio temático es una tarea que plantea el análisis de los requerimientos de usuario y el análisis de los requerimientos técnicos.

Requerimientos de usuarios:

- Ingreso ó control de los metadatos del documento que quiere incorporar al repositorio.
- Establecer mecanismos de seguridad para el contenido.
- Ofrecer un proceso estándar (depósito, revisión, publicación, preservación, búsqueda, navegación).

Requerimientos técnicos:

- Controlar los formatos de los archivos.
- Asegurar que se use el modelo de metadatos elegido.
- Elaborar un plan de preservación de documentos.
- Obtener la autorización escrita de los autores para conservar y dar acceso a largo plazo los contenidos depositados.
- Controlar el acceso de los usuarios.

Diseño

El diseño abarcará el diseño de la arquitectura del repositorio, la modalidad de publicación, el formato de los contenidos a archivar, la organización del contenido y la definición de quienes aportan los metadatos.

Para el diseño de la arquitectura del repositorio se recomienda la propuesta de los autores Alexander Ochoa Agüero, Athos Sánchez Mansolo, Rafael Jorge Pozo Lauzán [24], elegir un sistema centralizado, en el cual todos los contenidos se encuentran en una base de datos única.

Con respecto a la modalidad de publicación, el formato de los contenidos a archivar, la organización del contenido y la definición de quienes aportan los metadatos se deben considerar las políticas definidas.

Selección de tecnología e implementación

La selección de la tecnología requiere de un proceso de selección de las plataformas disponibles (software), las mismas se encuentran clasificadas de acuerdo al tipo de licenciamiento:

- Licencia Comercial: poseen la desventaja del alto costo de la licencia, código fuente cerrado, si se quiere integrar el repositorio con otros sistemas informáticos se cuenta con una relación de dependencia con un solo proveedor y se debe pagar por la implementación. Ej.: CONTENTdm [25], Digital Commons [26], DigiTool [27], EQUILLA [28], intraLibrary [29], Open Repository [30].
- Licencia Freeware: aunque es un software gratis, es de código fuente cerrado, no es multiplataforma. Ej.: Zentity de Microsoft [31]
- Licencia Open Source: posee el beneficio de compartir el código, no contempla restricciones de uso o modificación. Ej.: DSpace [32], E-Prints [33], Fedora [34], Greenstone [35], CDSware [36], Connexions [37] PlanetDR[38], DOOR [39], SciELO [40].

Estos paquetes de software se caracterizan por no necesitar de excesiva configuración o adaptación por parte de los administradores, aunque existen casos en los que se precisan de programas complementarios, por ejemplo, para la implementación de interfaces. [41]

Estos sistemas ofrecen una serie de funcionalidades básicas para el mantenimiento y gestión del repositorio, entre las que podemos mencionar: gestión del flujo documental (envío de documentos); evaluación, aceptación o rechazo de documentos); gestión de metadatos (incorporación, edición, revisión y evaluación de metadatos sobre los recursos); gestión de contenidos (facilitan la creación de colecciones de materiales y generalmente aceptan múltiples formatos de archivo); con respecto al uso del archivo existen diversas opciones de búsqueda y navegación, así como funciones de visualización y descarga.

En general, ofrecen interfaces de usuario, de autor y de administrador, y permiten la creación de diversos grupos de usuario y niveles de acceso.

Para realizar la selección de la tecnología se recomienda centrarse en los siguientes aspectos [42]:

- Interfaz: La forma de presentación al usuario final, así como la presentación a la persona que se ocupa del procesamiento.
- Flexibilidad: Adaptación de la herramienta, según las características institucionales.
- Lenguaje: Idiomas del ambiente de procesamiento y de la interfaz de recuperación.
- Contenidos: Formato de los documentos que acepta en sus colecciones.
- Procesamiento: Facilidades para procesar los documentos para una recuperación efectiva.
- Recuperación: Formas que tiene el usuario de acceder a los documentos.
- Requerimientos de sistema: Características de las computadoras que soportarán la herramienta y de las que harán uso de las colecciones.
- Servidor Web: Requerimientos de los servidores en los que se soportará la herramienta.
- Licencia: Si es libre o privada.
- Costo: Gratis o pago.

Conclusiones

Los artículos científicos, tesis doctorales, proyectos de investigación y otros tipos de documentos en la temática elegida para construir un repositorio requieren un modo de gestión que los describa y analice, los organice y lo presente de acuerdo con la filosofía del AA, para ello es preciso contar con directrices que ofrezcan una guía para el proceso de creación del repositorio.

Las directrices presentadas en este trabajo se estructuran en cuatro puntos referencia: finalidad, funciones, políticas y actividades, se ha considerado que todos los puntos definidos deben ser de cumplimiento básico para la creación de un repositorio temático.

Los resultados que se esperan lograr siguiendo las directrices expuestas son:

- Proporcionar un espacio donde almacenar información en la temática elegida.
- Preservar la producción científica de la temática seleccionada.
- Maximizar la visibilidad de la producción científica.
- Retroalimentar la investigación en el área temática.
- Integrar el repositorio en una plataforma tecnológica de fácil acceso.
- Crear y desarrollar comunidades electrónicas científicas.
- Facilitar la recuperación del contenido del repositorio desde la comunidad científica y el conjunto a través de la Web.

Con este trabajo se busca presentar los repositorios temáticos y una guía para su creación ya que existe poca información disponible sobre el desarrollo de los mismos en comparación con la literatura encontrada sobre repositorios institucionales.

Con respecto al AA, consideramos que a través de la vía verde se estimula el proceso de autoarchivo ofrecido por los repositorios temáticos, de esta manera se llega más rápidamente a investigadores interesados, beneficiando el área de conocimiento y favoreciendo el ciclo de investigación en una temática.

A modo de conclusión final, la importancia de desarrollar repositorios temáticos bajo las iniciativas del AA radica en el hecho de ser una fuente de información que cubre un área del conocimiento, en la cual los investigadores tienen contenidos actualizados sin ninguna barrera económica.

Referencias

- [1] Vázquez, G.; “Repositorios Digitales de Documentos”, Primer Curso sobre Escritura Científica y Repositorios de Acceso Abierto, 6 y 7 de diciembre de 2007
- [2] J. Arancibia, J. Santillán Aldana, I. Subirats Colls; Iniciativas de Acceso Abierto en Ciencias de la Información y Documentación: evolución y perspectivas de E-LIS, Revista Española de Documentación Científica, Vol. 28, N°2, pp 221-232. (2005)
- [3] S. Gonzalez Sanchez, “Repositorios de Información” (2007), [online] Disponible: <http://www.slideshare.net/salgonsan/tema-4-repositorios-de-informacin>, [Consulta: 1 de Noviembre de 2011]
- [4] S. Zambrano, “Repositorios de Información”, (2009) [online] Disponible: <http://www.slideshare.net/solanghyz/repositorios-de-informacin>, [Consulta: 1 de Noviembre de 2011]
- [5] Budapest Open Access Initiative. (2002) [online]. Disponible: <http://www.soros.org/openaccess/read.shtml> [Consulta: 18 de marzo de 2011].
- [6] Lynch, C; “Institutional repositories: essential infrastructure for scholarship in the digital age.” ARL No. 226: 1-7 (2003). [online] Disponible: http://scholarship.utm.edu/21/1/Lynch,_IRs.pdf [Consulta: 23 de Marzo de 2012]
- [7] PROJECT NECOBELAC. Red de Colaboración Entre Europa y Countries de América Latina y El Caribe. [online] Disponible: <http://62.204.194.27/necobelac/> [Consulta: 23 de Marzo de 2012]
- [8] Guimarães, M.; da Silva, C. H.; Horsth Noroña, I; Los repositorios temáticos en la estrategia de la iniciativa Open Access, Nutr Hosp 2012;27(Supl. 2):34-40, ISSN (Versión electrónica): 1699-5198

- [9] I. Daudinot Founier, Organización y recuperación de información en Internet: teoría de los metadatos, Acimed (2006); 14
- [10] Convención de Santa Fe, Historia y desarrollo del OAI-PMH [online] Disponible:<<http://travesia.mcu.es/portalnjb/jspui/html/10421/1823/page2.htm>> [Consulta: 1 de Noviembre de 2011]
- [11] Y. Sarduy Domínguez, P. Urra Gonzalez, Herramientas para la creación de colecciones digitales, Revista Acimed, Vol. 14 N° 5, (2006) [online] Disponible: <http://bvs.sld.cu/revistas/aci/vol14_5_06/aci19506.htm> [Consulta: 1 de Noviembre de 2011]
- [12] S. Weitzel, et all. “E-LIS: um repositório digital para a biblioteconomia e ciência da informação no Brasil.” XV Seminário Nacional de Bibliotecas Universitárias, São Paulo, Brasil, 10-14 de noviembre, (2008).
- [13] F. Calzada Prado, “Repositorios Digitales, bibliotecas digitales y CRAI”, Ed. Alfagrama, ISBN 978-987-1305-57-5, (2010)
- [14] Ley de Propiedad Intelectual, [online] Disponible: <http://www2.mincyt.gov.ar/11723.htm>, [Consulta: 1 de Noviembre de 2011]
- [15] Organización Mundial de la Propiedad Intelectual, ¿Qué es La propiedad intelectual?, OMPI N° 450(S). ISBN 92-805-1157-0
- [16] R. Feltrero Oreja, El Software Libre y la Producción y Socialización del Conocimiento en la Red: El problema de la Propiedad Intelectual. (2003)
- [17] Fundación Copyleft, [online] Disponible:< <http://copyleft.org/> > [Consulta: 23 de Marzo de 2012]
- [18] Licencias de Documentación, Disponible: http://www.reeducativa.com/webquests/licencias/licencias_documentacion.html. [Consulta: 23 de Marzo de 2012]
- [19] I. Subirats Coll, Principales factores que pueden dificultar la evolución de un archivo abierto y cómo afrontarlos, Boletín de la SEDIC N° 42, ISSN: 1137-0904 [online] Disponible:<http://www.sedic.es/p_boletinclip42_confirmar_imprimir.htm> [Consulta: 1 de Noviembre de 2011]
- [20] G. Bueno de la Fuente, “Modelo de repositorio institucional de contenido educativo (RICE): la gestión de materiales digitales de docencia y aprendizaje en la biblioteca universitaria”, Tesis Doctoral, Universidad Carlos III De Madrid, (2010), [online] Disponible: http://e-archivo.uc3m.es/bitstream/10016/9154/1/Tesis%20doctoral-Gema_Bueno.pdf, [Consulta: 18 de marzo de 2011]
- [21] M. Doria, C. Inchaurredo, “Iniciativas del Acceso Abierto a la información científica y académica en el campo de la ingeniería”, Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación N°7, ISSN 1850-9959, Junio 2012, pp 66-75]
- [22] Tesouro UNESCO, Disponible: <http://databases.unesco.org/thessp/>, [Consulta: 23 de Marzo de 2012]
- [23] IEEE Standard glossary of software engineering terminology (1997)
- [24] A. Ochoa Agüero, A. Sanchez Manzolo, R. Pozo Lauzán, Repositorios De Objetos De Aprendizaje De Acceso Abierto Para La Educación De Postgrado. República Bolivariana De Venezuela. 2010 [online] Disponible: <http://www.rcim.sld.cu/revista_21/articulo_pdf/repositorioaprendizaje.pdf> [Consulta: 23 de Marzo de 2012]
- [25] CONTENTdm [online] Disponible: www.CONTENTdm.org [Consulta: 23 de Marzo de 2012]
- [26] Digital Commons [online] Disponible: <http://digitalcommons.bepress.com> [Consulta: 23 de Marzo de 2012]
- [27] DigiTool [online] Disponible: <http://www.exlibrisgroup.com/category/DigiToolOverview> [Consulta: 23 de Marzo de 2012]
- [28] EQUILLA [online] Disponible: www.equella.com [Consulta: 23 de Marzo de 2012]
- [29] IntraLibrary [online] Disponible: www.intrallect.com [Consulta: 23 de Marzo de 2012]
- [30] Open Repository [online] Disponible: www.openrepository.com [Consulta: 23 de Marzo de 2012]
- [31] Zentity de Microsoft [online] Disponible: research.microsoft.com/en-us/projects/zentity/ [Consulta: 23 de Marzo de 2012]
- [32] DSpace [online] Disponible: www.dspace.org [Consulta: 23 de Marzo de 2012]
- [33] E-Prints [online] Disponible: <http://www.eprints.org/> [Consulta: 23 de Marzo de 2012]
- [34] Fedora [online] Disponible: <http://fedoraproject.org/es/> [Consulta: 23 de Marzo de 2012]

- [35] Greenstone [online] Disponible: <http://www.greenstone.org/> [Consulta: 23 de Marzo de 2012]
- [36] CDSware [online] Disponible: <http://invenio-software.org/> [Consulta: 23 de Marzo de 2012]
- [37] Connexions [online] Disponible: <http://cnx.org/> [Consulta: 23 de Marzo de 2012]
- [38] PlanetDR [online] Disponible: <http://planet.urv.es/planetdr/> [Consulta: 23 de Marzo de 2012]
- [39] DOOR [online] Disponible: www.door.sourceforge.net/ [Consulta: 23 de Marzo de 2012]
- [40] SciELO [online] Disponible: <http://www.scielo.org> [Consulta: 23 de Marzo de 2012]
- [41] G. Bueno de la Fuente, D. Rodríguez Mateos, "Herramientas de software para OAIPMH", La Iniciativa de Archivos Abiertos (OAI): situación y perspectivas en España y Latinoamérica. Bogotá: Rojas Eberhard, 2007, pp. 247-302. ISBN 978-958-9121-89-4
- [42] Y. Sarduy Domínguez, P. Urra González, Herramientas para la creación de colecciones digitales, Revista Acimed, Vol. 14 N° 5, (2006) [online] Disponible: http://bvs.sld.cu/revistas/aci/vol14_5_06/aci19506.htm [Consulta: 1 de Noviembre de 2011]

Claudia Inés Inchaurredo. Mg. en Ingeniería de Software. Ing. en Sistemas de Información. Profesor Titular Dedicación Semiexclusiva. Cátedra: Sistemas de Gestión. Fac. de Tecnología y Cs. Aplicadas de la UNCa. Cat. IV Prog. De Incentivos.

Germán Antonio Montejano. Mg. en Ingeniería de Software. Profesor UNSL y UNLPam. Categoría I Programa de Incentivos. Director de proyectos de investigación acreditado incentivado UNLPam y UNSL. Codirector carrera Maestría en Calidad del Software y Especialización Ing. Software UNSL.

Dirección de Contacto del Autor/es:

Maria Vanesa Doria
 Maximio Victoria 55 - C.P: 4700
 San Fernando del Valle de Catamarca
 Argentina
 e-mail: vanesadoria@gmail.com

Claudia Inés Inchaurredo
 Maximio Victoria 55 - C.P: 4700
 San Fernando del Valle de Catamarca
 Argentina
 e-mail: cinchaurredo@sistemas.frc.utn.edu.ar

Germán Antonio Montejano
 Ejército de los Andes 950 – CP: 5700
 San Luis
 Argentina
 e-mail: gmonte@unsl.edu.ar

Maria Vanesa Doria. Lic. en Sistemas de Información. Jefe de Trabajos Regular Dedicación Semiexclusiva. Cátedras: Ingeniería de Software III y Sistemas de Gestión Fac. de Tecnología y Cs. Aplicadas de la UNCa. Cat. IV Prog. De Incentivos