

# Propuesta Metodológica para el desarrollo de un Repositorio de Acceso Abierto

María V. Doria <sup>1</sup>, Germán A. Montejano <sup>2,3</sup> & Claudia I. Inchaurredo <sup>1</sup>

1 Dpto. Informática – Facultad de Tecnología y Cs. Aplicadas – Universidad Nacional de Catamarca

Maximio Victoria 55 - C.P: 4700 - San Fdo. del Valle de Catamarca- TEL. 03833- 435112  
vanesadoria@gmail.com, cinchaurrondo@sistemas.frc.utn.edu.ar

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  
gmonte@unsl.edu.ar

## Resumen

En el presente trabajo se expone la metodología empleada para el desarrollo de un repositorio de acceso abierto (AA).

A través de esta propuesta metodológica se pretende brindar las actividades a seguir para el desarrollo, funcionamiento y utilización de un repositorio. Siendo un tema de actualidad y teniendo un uso creciente en diferentes ámbitos se consideró adecuado abordar el tema desde la ingeniería de software.

La Metodología se desarrolla en combinación con la inteligencia de negocio por medio de una de las variaciones de la minería de datos llamada minería de texto que permite la recuperación de información.

La literatura existente sobre desarrollo de repositorios desde la perspectiva de la ingeniería de software es escasa comparada con la ingente información sobre repositorios desde el campo de la bibliotecología y las ciencias de la documentación.

**Palabras clave:** repositorios, acceso abierto (AA), metodología para el Análisis, Diseño y Desarrollo de repositorios.

## Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: Ingeniería de Software, Conceptos, Métodos y Herramientas en un Contexto de “Ingeniería de Software en Evolución” – Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

## Introducción

Durante los últimos años en Internet han proliferado los repositorios de información como una alternativa a los métodos tradicionales de publicación y difusión de la producción científica y académica.

Por consiguiente, se hace necesario saber qué se entiende por repositorios de información para proceder a una adecuada planificación, diseño y desarrollo de un repositorio de acceso abierto (AA).

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.

Entonces los repositorios de AA son un lugar donde se guardan documentos científicos y/o académicos siguiendo la filosofía del AA. Esta filosofía manifiesta 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” [1].

A modo orientativo, en un trabajo anterior [2] se presentó el siguiente flujo de actividades básico para el desarrollo de un repositorio de AA:

- Análisis documental del cuerpo de documentos que formaran parte del repositorio.
- Análisis de requerimientos.
- Diseño.
- Selección de tecnología e implementación.

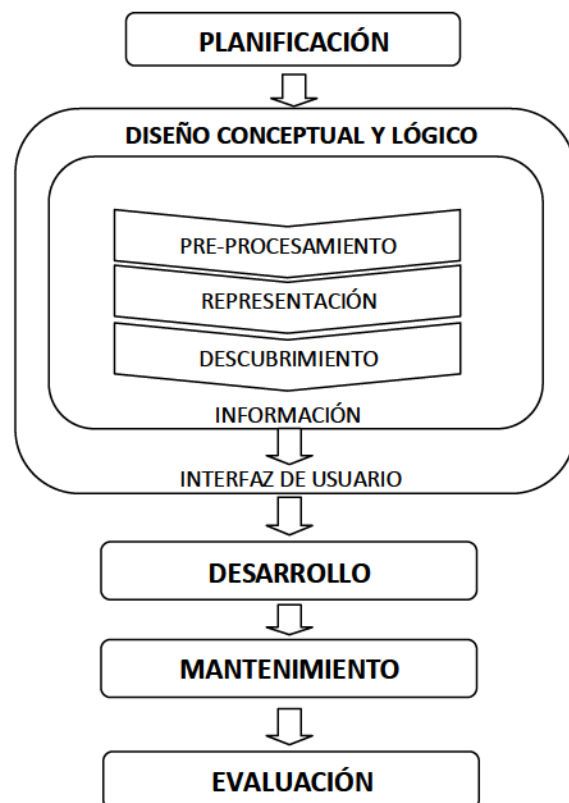
Teniendo en cuenta este flujo básico se consideró el uso de la “Metodología para el Análisis, Diseño y Desarrollo de

Servicios de Información Digital” propuesta por Navarro [3] ya que ella considera que los repositorios de información pertenecen a la familia de los servicios digitales y se integran las técnicas de la inteligencia de negocio por medio de una de las variaciones de la minería de datos llamada minería de texto [4] que facilita la recuperación de información documental.

Esta metodología establece cinco etapas:

- Etapa de planificación.
- Etapa de diseño conceptual y lógico.
- Etapa de desarrollo.
- Etapa de mantenimiento.
- Etapa de evaluación.

En la Figura 1 se muestra en qué etapa se ha combinado la metodología elegida con la minería de texto.



**Figura 1.** Metodología para el Análisis, Diseño y Desarrollo de Servicios de Información Digital combinada con Minería de Texto.

Cada una de las etapas que muestra la Figura 1 tiene un fin distinto y se lleva a cabo mediante una serie de actividades y tareas concretas que producen resultados.

El paso de una etapa a otra no necesariamente es lineal, habitualmente puede ser conveniente retroceder a la etapa anterior para replantear alguna conclusión adoptada.

A continuación se describen brevemente las etapas de esta metodología [3]:

- Etapa de Planificación: consiste en la determinación de un conjunto de elementos clave que regulan y guían todo el ciclo de diseño y desarrollo del repositorio de AA. Consta de las siguientes actividades:
  1. Determinación de la Finalidad, Funciones y Comunidad de usuarios del repositorio.
  2. Expresar con enunciados generales los servicios que ofrecerá el repositorio.
  3. Definición de los objetivos generales se trata de describir los resultados concretos que se deben lograr con el repositorio
  4. Definición de objetivos específicos a partir de los objetivos generales, deben ser muy concretos, realistas y complementarios entre sí.
  5. Previsión de los recursos humanos.
- Etapa de Diseño Conceptual y Lógico: consiste en una presentación detallada de los principales aspectos del proyecto y está conformada por dos momentos. Primero se realiza un *diseño conceptual*, es la representación y descripción de lo que se pretende construir independiente de cualquier condicionamiento tecnológico y luego un *diseño lógico*, adapta el diseño conceptual con las

prescripciones y los condicionantes de la realidad.

Las actividades del Diseño Conceptual son:

1. Indicación y análisis de los requerimientos operacionales del repositorio.
2. Definición de entradas al sistema y localización de las fuentes de información adecuadas.
3. Establecimiento de los procesos básicos para el tratamiento de los documentos. En esta actividad se integra las tres técnicas clásicas de minería de texto [4]:(1) Pre-procesamiento, es el proceso mediante el cual los textos se transforman en algún tipo de representación estructurada que facilite su análisis; (2) Representación, depende de la técnica de pre-procesamiento utilizada y determinará el algoritmo de descubrimiento a utilizar y (3) Descubrimiento, son algoritmos que, a partir de una representación estructurada de la información, son capaces de descubrir regularidades en los textos.
4. Definición de Políticas.
5. Descripción de la arquitectura e interfaz del sistema.

Las actividades propias del Diseño Lógico son:

1. Decisión de las herramientas informáticas.
2. Traducción de los requerimientos funcionales del modelo diseñado en especificaciones de requisitos técnicos.
3. Transformación del modelo conceptual a los modelos de prestaciones y de datos que proporcionan las aplicaciones informáticas seleccionadas.

- Etapa de Desarrollo: es el cumplimiento de las actividades planificadas para la creación del repositorio y su puesta en marcha. Las actividades de la etapa son:
  1. Obtención e instalación del software Greenstone.
  2. Creación del prototipo.
  3. Adquisición y carga de recursos informáticos.
  4. Control del funcionamiento del prototipo.
  5. Prueba del prototipo.
  6. Redacción de manual de usuario.
- Etapa de mantenimiento: Cuando el repositorio esté funcionando en su versión definitiva se deben realizar trabajos de mantenimiento de los recursos de información y de los servicios que se ofrecen, estas tareas son:
  1. Establecer procedimientos para recibir, registrar y hacer seguimiento de los problemas y peticiones de modificación de los usuarios para determinar tipo de mantenimiento que se debe realizar, alcance y criticidad.
  2. El administrador técnico y el administrador de contenido deberá reproducir o comprobar el problema.
  3. Revisión/aceptación del mantenimiento.
- Etapa de Evaluación: consiste en tres acciones sucesivas y complementarias:
  1. Medición: a partir de datos recogidos a través de una encuesta cerrada a expertos y potenciales usuarios con el fin de determinar si el repositorio ha logrado los objetivos fijados.
  2. Emisión: expresar un juicio de valor a partir de los resultados de la medición, comparando el

funcionamiento real del repositorio con los resultados esperados.

3. Determinar recomendaciones y propuestas de solución a las dificultades encontradas en el repositorio en las acciones de medición y juicio.

## **Líneas de Investigación**

Se plantean diferentes líneas de investigación:

En una primera instancia se examina el diseño y desarrollo de repositorios de información tanto institucionales como temáticos.

En una segunda instancia se analizan las metodologías empleadas para la construcción de repositorios de información y las técnicas de Minería de Texto para la recuperación de la información.

En una tercera instancia se analizan y comparan herramientas de software para el desarrollo de repositorios de información.

Avanzar en las líneas de investigación permitió presentar esta propuesta de metodología para el desarrollo de repositorios de AA.

## **Resultados y Objetivos**

Los resultados obtenidos hasta el momento son:

- Se logró combinar la minería de texto para la recuperación de documentos de información en la etapa de diseño de la metodología propuesta.
- Se concretaron las actividades de la etapa de planificación, diseño conceptual y lógico, desarrollo y mantenimiento.

Siguiendo la propuesta metodológica para el desarrollo de un repositorio de acceso

abierto se construyó el Repositorio en Ingeniería de Software (RIS).

Actualmente se están estudiando los recursos para poner el repositorio en línea y luego realizar la etapa de evaluación.

Como trabajo futuro se llevará a cabo la etapa de evaluación para ello el repositorio desarrollado estará disponible en la web.

Los objetivos planteados para llegar a los resultados mencionados son:

- Describir la metodología que permita el desarrollo del prototipo de Repositorio Temático de AA.
- Precisar los pasos para el desarrollo del repositorio de AA con la plataforma de software libre seleccionada.

## Formación de Recursos Humanos

Esta investigación se desarrolla dentro del proyecto mencionado en el contexto y corresponde a una Tesis de Maestría en curso titulada “Modelo de un Repositorio en Ingeniería de Software para la Facultad de Tecnología y Cs. Aplicadas de la UNCa.” de la carrera de postgrado “Maestría en Ingeniería de Software” de la Universidad Nacional de San Luis.

La tesis fue motivo de una postulación y adjudicación de la beca PROFITE (Programa de Becas para la Finalización de Tesis de Posgrado para docentes de Universidades Nacionales), la misma es financiada por el Ministerio de Educación, Subsecretaría de Gestión y Coordinación de Políticas Universitarias, con el proyecto denominado: “Construcción de un Repositorio Temático de Acceso Abierto con Software Libre en el ámbito universitario”.

El equipo de trabajo está formado por:

- Tesista: Lic. Maria Vanesa Doria.
- Director: Doctor Germán Antonio Montejano.

- Co – Director: Mg. Ing. Claudia Inés Inchaurredo

De acuerdo a la línea de investigación de la tesis, se participó en la capacitación “Repositorios digitales institucionales: Diseño, implementación y optimización de un recurso estratégico para las Universidades.” Brindado en el marco del “Proyecto Investigación y desarrollo en repositorios institucionales: aplicaciones y experiencias en universidades de la región bonaerense.” (PICT-O-CIN 2010-149).

Se prevé asimismo continuar con la formación y fortalecimiento en investigación de la tesista a través de la guía del director y codirector.

## Referencias

- [1] Budapest Open Access Initiative. (2002) [online]. Disponible: <<http://www.soros.org/openaccess/read.shtml>> [Consulta: 18 de marzo de 2011]
- [2] Doria, M; Inchaurredo C. & Montejano G. Directrices para la construcción de un repositorio temático, Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología, pag. 40-49 ISSN 1850-9959, Abril 2013
- [3] Navarro, E., Planificación, diseño y desarrollo de servicios de información digital. En: Tramullas J, Garrido P (eds.). Software libre para servicios de información digital. Madrid: Pearson Prentice-Hall; ISBN 978-84-8322-299-7, 2006. p. 23-43
- [4] Botta-Ferret, E & Cabrera-Gato J.; Minería de textos: una herramienta útil para mejorar la gestión del bibliotecario en el entorno digital; ACIMED v.16 n.4 2007, ISSN 1024-9435.