

UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO



FACULTAD DE INGENIERÍA

**ACERVO DIGITAL DE RECURSOS
DIDÁCTICOS**

TESIS

PARA OBTENER EL TÍTULO DE
INGENIERO EN COMPUTACIÓN

PRESENTA:
RODRIGO ENRIQUE DÍAZ ACAL

DIRECTORES DE TESIS:
M.I. MAURICIO MORGADO CASTILLO
ING. LAURA SANDOVAL MONTAÑO



DICIEMBRE 2012

CIUDAD UNIVERSITARIA, MÉXICO, D.F.

ÍNDICE

1. Introducción.....	9
2. Objetivo.....	13
3. Marco teórico.....	15
3.1 Educación.....	16
3.1.1 Enseñanza.....	16
3.1.2 Aprendizaje.....	16
3.1.3 Procesos para la generación y publicación de contenido.....	18
3.1.4 Revisión por expertos.....	21
3.2 Procesos de gestión de información.....	24
3.3 Tecnologías involucradas.....	31
3.3.1 Formatos de archivo.....	31
3.3.2 Acervos digitales.....	34
3.4 Bases de datos.....	37
3.4.1 Sistema de gestión de bases de datos (DBMS).....	40
3.5 Metadatos.....	43
3.5.1 Modelos de metadatos.....	45
3.6 Web 2.0.....	49
4. Contexto y análisis del problema.....	51
4.1 Situación actual del Departamento de Portugués en el CELE.....	52
4.2 Uso e impacto de los acervos digitales en la educación.....	54
4.3 Análisis de acervos implementados en instituciones educativas.....	65
4.3.1 MIT OpenCourseWare.....	66
4.3.2 Multimedia Educational Resource for Learning and Online Teaching.....	68
4.3.3 Centro de Recursos para la Enseñanza y el Aprendizaje.....	71
4.3.4 Desarrolla, Aprende y Reutiliza.....	73
4.3.5 Funes: repositorio digital de documentos en Educación Matemática.....	75

5. Propuesta.....	79
5.1 Resultados de la comparativa entre los acervos seleccionados.....	80
5.2 Enfoque	86
5.2.1 Proceso de Gestión del Proyecto.....	87
5.2.2 Proceso de Implementación del Software.....	90
5.3 Diseño.....	93
6. Implementación.....	101
6.1 Arquitectura.....	102
6.1.1 Modelo de datos	104
6.1.2 Despliegue de la Didactiteca de Portugués.....	107
6.2 Desarrollo técnico	109
6.2.1 Integración con RU-CELE.....	110
6.2.2 Pruebas de rendimiento.....	111
6.3 Resultados	124
7. Conclusiones.....	127
8. Trabajo a futuro.....	133
8.1 Configuración flexible.....	134
8.2 Visualización de materiales en el sitio	135
8.3 Elementos de difusión	138
9. Apéndices.....	143
9.1 Control de versiones.....	144
10. Bibliografía	147

ÍNDICE DE FIGURAS

<i>Figura 1 - Proceso para la generación de contenido</i>	<i>19</i>
<i>Figura 2 - Proceso de revisión</i>	<i>22</i>
<i>Figura 3 - Ciclo de gestión de información</i>	<i>25</i>
<i>Figura 4 - Procesos previos al almacenamiento.....</i>	<i>27</i>
<i>Figura 5 - Formatos de imagen</i>	<i>32</i>
<i>Figura 6 - Sistema de base de datos</i>	<i>38</i>
<i>Figura 7 – Ejemplo de metadatos</i>	<i>43</i>
<i>Figura 8 - Metadatos de objetos.....</i>	<i>45</i>
<i>Figura 9 - CMS a nivel mundial</i>	<i>48</i>
<i>Figura 10 – Conceptos de Web 2.0.....</i>	<i>49</i>
<i>Figura 11 - Contenidos predominantes.....</i>	<i>55</i>
<i>Figura 12 - Fuentes de información para decidir participar en un repositorio</i>	<i>57</i>
<i>Figura 13 - Línea del tiempo OpenDOAR</i>	<i>58</i>
<i>Figura 14 - Repositorios en Norteamérica</i>	<i>59</i>
<i>Figura 15 - Tipos de contenido en México</i>	<i>60</i>
<i>Figura 16 - Uso de CMS en México</i>	<i>82</i>
<i>Figura 17 - CMS en México agrupando los desarrollos propios.....</i>	<i>83</i>
<i>Figura 18 - Procesos básicos</i>	<i>87</i>
<i>Figura 19 - Proceso de Gestión del Proyecto.....</i>	<i>89</i>
<i>Figura 20 - Proceso de Implementación del Software</i>	<i>92</i>
<i>Figura 21 - Diagrama de casos de uso de la Didactiteca</i>	<i>95</i>
<i>Figura 22 - Diagrama de paquetes</i>	<i>96</i>
<i>Figura 23 - Diagrama ER de la Didactiteca de Portugués</i>	<i>99</i>
<i>Figura 24 - Modelo MVC</i>	<i>103</i>
<i>Figura 25 - Modelo relacional</i>	<i>105</i>
<i>Figura 26 - Modelo relacional</i>	<i>106</i>
<i>Figura 27 - Diagrama de despliegue de la Didactiteca de Portugués.....</i>	<i>108</i>
<i>Figura 28 - Tiempos de respuesta: 16 usuarios.....</i>	<i>115</i>
<i>Figura 29 - Tiempo de respuesta promedio: 16 usuarios.....</i>	<i>116</i>
<i>Figura 30 - Tiempos de respuesta: 300 usuarios.....</i>	<i>118</i>
<i>Figura 31 - Tiempo de respuesta promedio: 300 usuarios.....</i>	<i>118</i>
<i>Figura 32 - Tiempos de respuesta: prueba de estrés</i>	<i>120</i>
<i>Figura 33 - Tiempo de respuesta promedio: prueba de estrés</i>	<i>122</i>
<i>Figura 34 - Reproducción de video en el MIT OCW.....</i>	<i>136</i>
<i>Figura 35 - MIT OCW: enviar por email</i>	<i>139</i>
<i>Figura 36 - MIT OCW: forma para envío de email</i>	<i>139</i>
<i>Figura 37 - CREA: compartir por Twitter.....</i>	<i>140</i>
<i>Figura 38 - CREA: redactar mensaje en Twitter</i>	<i>140</i>
<i>Figura 39 - MERLOT: código QR.....</i>	<i>141</i>
<i>Figura 40 - Control de versiones centralizado</i>	<i>144</i>
<i>Figura 41 - Control de versiones distribuido</i>	<i>145</i>

ÍNDICE DE TABLAS

<i>Tabla 1 - Matriz de Estructuración.....</i>	<i>29</i>
<i>Tabla 2 - Ejemplos de números mágicos.....</i>	<i>34</i>
<i>Tabla 3 - Repositorios en México</i>	<i>61</i>
<i>Tabla 4 - CMS en los acervos analizados.....</i>	<i>81</i>
<i>Tabla 5 - Tipos de materiales y sus formatos.....</i>	<i>84</i>
<i>Tabla 6 - Esquemas de metadatos</i>	<i>85</i>
<i>Tabla 7 - Tipos de búsqueda</i>	<i>85</i>
<i>Tabla 8 - Tiempos de respuesta de acervos y repositorios.....</i>	<i>113</i>
<i>Tabla 9 - Prueba de rendimiento: 16 usuarios</i>	<i>115</i>
<i>Tabla 10 - Prueba de rendimiento: 300 usuarios</i>	<i>117</i>
<i>Tabla 11 - Prueba de estrés: resultados.....</i>	<i>119</i>
<i>Tabla 12 - Prueba de estrés: detalles</i>	<i>121</i>
<i>Tabla 13 - Pruebas de rendimiento: búsquedas.....</i>	<i>123</i>

AGRADECIMIENTOS

A mis padres, por guiarme en el camino que me ha llevado hasta este momento. Su apoyo y consejos han sido fundamentales en mi crecimiento personal y profesional.

A mis profesores, quienes me armaron con conocimiento para enfrentar problemas presentes y venideros.

A mi universidad, que me ofrece tantos lugares y oportunidades. Aulas, laboratorios y bibliotecas para todo lo que mis estudios requieren. Conciertos, obras, museos y áreas verdes para mi diversión y enriquecimiento cultural.

A mis directores de tesis, por ayudarme en la creación del presente trabajo.

INTRODUCCIÓN

Durante años asistí a clases de inglés, he estado presente en algunas clases de alemán en el Centro de Enseñanza de Lenguas Extranjeras (CELE), he sabido de las clases de italiano, francés y japonés que se dan también dentro de este mismo plantel educativo, gracias a la experiencia de algunos amigos.

Cada idioma se aprende de forma distinta, pero todas las clases que conozco hacen uso de un recurso distintivo: el libro de ejercicios. Usualmente viene acompañado de un libro de texto. Ambos forman parte de una colección especial, diseñada para separar el aprendizaje de un idioma por niveles. Conforme un alumno avanza de nivel cambia de libro, aunque en algunos casos varios niveles pueden compartir el mismo libro.

He notado que frecuentemente los profesores encuentran insuficiente el guiarse únicamente con un libro de ejercicios, tal vez porque necesitan enseñar a sus alumnos algo que no viene contemplado en tales ejemplares, como factores culturales y de actualidad, aunque también puede ser simplemente la necesidad de hacer hincapié en algún tópico específico. Cualquiera que sea el caso, es común que un profesor acuda a sus propios métodos de enseñanza, ayudado por fragmentos de texto literarios, imágenes, películas, música, juegos, etc., materiales que cada profesor recopila e integra a sus clases.

En el Centro de Enseñanza de Lenguas Extranjeras, el Departamento de Portugués propone una actualización curricular, la cual implica el diseño de un nuevo plan de estudios y los programas correspondientes para la enseñanza de la lengua portuguesa. Dentro de esta propuesta se definen un conjunto de tareas, las cuales están diseñadas para ser utilizadas por los alumnos desde su ingreso hasta la conclusión de sus estudios.

La actualización curricular requiere de una cantidad de recursos educativos que sea capaz de abarcar la totalidad de tareas planteadas. Tales recursos no serán adquiridos como lo puede ser un libro de ejercicios, sino que serán creados, en un inicio, por los profesores del Departamento de Portugués. Cada profesor tiene un modo distintivo de enseñar, aun cuando se valga de los mismos medios que sus coetáneos. Es en las actividades creadas por los profesores, donde se puede tener una distinción más clara de la forma en que una persona transmite su conocimiento. Por tanto, al contar con un conjunto de recursos creados por los profesores, se da pie a una mayor diversidad en cuanto a la enseñanza del idioma portugués.

Para cumplir con su meta, el Departamento de Portugués requiere un lugar en el cual los profesores puedan colocar los recursos educativos que generan. Un lugar en donde puedan almacenar, además, los materiales que una actividad en clase requiere. Debe ser de fácil acceso, pero también capaz de resguardar tal producción intelectual de quien no forme parte de este proyecto.

Por ahora, y como se detalla en el capítulo **4.1**, el Departamento de Portugués cuenta únicamente con el Curso Activo de Portugués, cuyos recursos no han sido actualizados lo suficiente como para incluir temas actuales y de matiz cultural fresco; es decir, temas de interés para todo alumno que desee aprender el idioma.

Es en este punto donde la ingeniería es capaz de auxiliar en la enseñanza de la lengua portuguesa, al proveer el lugar que se necesita para colocar los recursos educativos que se requieran en el nuevo plan curricular.

Será un espacio virtual, un sistema de fácil acceso que un profesor podrá visitar desde la comodidad de su hogar. Las ventajas que ofrece la tecnología permitirán que el almacenamiento, estructuración, distribución y mantenimiento los recursos educativos puedan llevarse a cabo de manera eficiente.

Seguramente sonará familiar el uso de la computación en un entorno educativo. En general, suele ligarse al concepto de educación a distancia, donde los estudiantes no requieren asistir físicamente a un aula, al recibir toda la información y ser evaluados mediante herramientas que hacen uso de Internet. También es común pensar en cualquier otra clase de recurso destinada a los aprendientes, tanto si forman parte de un sistema escolarizado como si son independientes.

Pero también puede crearse un entorno destinado a los profesores, en el que puedan compartir actividades y materiales, retroalimentándose con las opiniones de sus colegas. Donde un grupo se encarga de verificar la calidad de los recursos didácticos, mientras otro vigila y administra.

Dentro de este entorno virtual será más fácil evitar que las actividades caigan en desuso o que se conserven temas que hayan dejado de interesar a los alumnos. Cada profesor será libre de elegir los recursos que crea convenientes y podrá buscarlos por medio de criterios que se adapten a sus necesidades (temática, tiempo que dura una actividad en clase, variante del idioma que se emplea, etc.).

Crear tal entorno es, precisamente, el tema de esta tesis.

OBJETIVO

Crear un sistema en donde los profesores del Departamento de Portugués puedan publicar y consultar diversas actividades para sus cursos presenciales en el Centro de Enseñanza de Lenguas Extranjeras (CELE). El sistema estará estrechamente ligado con la actualización curricular que ha tenido el Departamento de Portugués, donde se reestructuró el plan de estudios y enfoque de enseñanza.

- Facilitará la preservación e intercambio de recursos didácticos entre los profesores, puesto que serán ellos quienes contribuyan al crecimiento del sistema con sus aportaciones.
- Garantizará la calidad de las aportaciones al integrar un proceso de dictaminación, el cual estará completamente integrado en el sistema y facilitará el proceso revisión.
- Será un espacio para el intercambio de opiniones, puesto que cada recurso podrá ser comentado y calificado por los usuarios.
- Incluirá procesos de gestión, para que un grupo de docentes pueda administrar tanto los recursos como los usuarios del sistema.

La creación del sistema debe contemplar una posible expansión del mismo, fuera del Departamento de Portugués, ya que posteriormente se pondrá a consideración compartir el material que contenga con otras instancias de la UNAM e instituciones externas.

MARCO TEÓRICO

3.1 Educación

3.1.1 Enseñanza

El concepto de enseñanza está estrechamente relacionado con el concepto de aprendizaje. Enseñar y aprender son dos términos fuertemente relacionados, apoyados entre sí por una relación dinámica (1). El aprendizaje es la adquisición de algo, la enseñanza implica dar algo.

Debido a que el término de aprendizaje puede expresarse en un área como el resultado de la enseñanza, es fácil mezclar ambos y decir que la tarea de la enseñanza es producir un aprendizaje productivo, cuando en realidad tiene más sentido expresar que *«una tarea central de la enseñanza es permitir al estudiante realizar las tareas del aprendizaje»* (2).

El presente trabajo no pretende incluir un análisis minucioso acerca de los modelos de enseñanza ni forjar una definición propia acerca de qué es enseñar. Sin embargo, podemos considerar a la enseñanza desde el punto de vista del docente o la institución educativa, y definirla como *«las estrategias que adopta la escuela para cumplir con su responsabilidad. Enseñanza no equivale meramente a instrucción, sino a la promoción sistemática del aprendizaje mediante diversos medios»* (3).

Esta acepción de la enseñanza es una de las más importantes porque supone atender de manera simultánea la práctica profesional, el plan de estudios y la organización de la clase. Para ello es conveniente, en primer lugar, humanizar el entorno de aprendizaje aplicando formas naturales de aprendizaje y examinar cómo se lleva a cabo la interacción entre la enseñanza y el aprendizaje. En segundo lugar, introducir una enseñanza dinámica de manera que el aprendizaje se produzca a través del plan de estudios, donde el aprendiz esté activamente comprometido intelectual, física y emocionalmente (2).

El binomio enseñanza-aprendizaje se basa en la interacción entre iguales, la negociación, el intercambio de experiencias y en la participación crítica (y activa) en espacios apropiados para la comunicación. Con ello el proceso de construcción y adquisición del conocimiento adquiere tanta importancia como los resultados del aprendizaje (4). De este modo, la enseñanza puede verse como la formación específica que proporciona la escuela, la aportación que la escuela realiza en el desarrollo intelectual de una persona.

3.1.2 Aprendizaje

El aprendizaje es una tarea que va intrínsecamente asociada con el desarrollo humano y consiste en *«un cambio de la disposición o capacidad humana, con carácter de relativa permanencia y que no es atribuible simplemente al proceso de desarrollo»* (5).

Asimismo, el profesor en el proceso de enseñanza-aprendizaje es quien pretende que la persona logre un aprendizaje. Entre sus funciones, se encuentra el diseñar, emprender y evaluar situa-

ciones de aprendizaje para que el alumno alcance logros específicos. Esta relación enseñanza-aprendizaje puede darse tanto en un proceso educativo como en uno instruccional. El aprendizaje se auxilia de la planificación y una adecuada selección de los medios a utilizar (1). Hoy en día, estos medios de enseñanza pueden incluir vídeos, televisión, conferencias, videoconferencias, libros de texto o presentaciones informáticas.

Es deseable que los profesores de todos los niveles educativos tengan la capacidad para promover en sus alumnos el aprendizaje de esos conocimientos. Se debe considerar la naturaleza del aprendizaje para poder proponer medios de enseñanza eficaces que produzcan aprendizajes significativos. Así, al conocer los procesos internos que llevan al alumno a aprender, manejará algunas ideas para propiciar, facilitar o acelerar el aprendizaje (4).

Para que el profesor pueda realizar mejor su trabajo, algunos autores (1) (3) consideran la reflexión, no sólo de su desempeño como docente, sino de la forma en que aprende el alumno, en cuáles son los procesos internos que lo llevan a aprender en forma significativa y en qué puede hacer para propiciar este aprendizaje.

Entonces, puede observarse que el trabajo del profesor incluye hacer posible el aprendizaje de sus alumnos, elaborando dinámicas y situaciones en las que pueda darse el proceso de aprender. Los alumnos adquieren muchos conocimientos fuera del salón de clases de manera cotidiana, pero dentro es donde aprenden lo que intencionalmente quiere enseñarles el profesor, aquellos conocimientos que considera vitales para su formación académica. El reto del profesor será lograr que los alumnos sean capaces de darle sentido a su conocimiento para que pueda ser utilizado para sus propios fines, y no sólo para fines escolares (4).

Una forma de aprendizaje es la inclusión de las Tecnologías de la Información y la Comunicación (TIC), tanto en el aprendizaje tradicional como en el aprendizaje a distancia. Las TIC en la educación pueden fungir como un medio o un recurso más para la educación, pero, desde un punto de vista pedagógico, la interconexión que crean no es necesariamente el principal recurso para el aprendizaje, pero sí lo es la calidad de dichas interacciones que puedan desarrollarse en función de los medios y recursos tecnológicos e informáticos.

Es importante considerar que las perspectivas para la educación en línea estarían tendientes a generar espacios para el conocimiento y formación del alumnado a través de diversos métodos. Pero es la práctica docente uno de los elementos decisivos para hacer propuestas innovadoras en el aprendizaje a partir de los nuevos modelos pedagógicos y del uso de las tecnologías, siendo la colaboración un factor considerable para rebasar las condiciones que pudieran limitar el proceso de aprendizaje en los ambientes virtuales (6).

Ante este panorama, todos aquellos involucrados (educadores, ingenieros, directivos, etc.) comparten, en mayor o menor medida, la responsabilidad para pensar, diseñar y evaluar juntos la introducción de nuevas tecnologías no solamente desde su aplicación educativa sino también

desde su función comunicativa (por ejemplo, para la difusión cultural y promoción). Esto incluye la educación a distancia y la educación presencial.

El aprendizaje se da en la medida en que el individuo se siente involucrado y en este sentido es que el ambiente mediado por tecnologías provoca procesos de aprendizaje, no es la tecnología en sí, sino su uso didáctico (4).

La aplicación de las TIC en la enseñanza propicia diversos cambios y definiciones (6), las más importantes a considerar incluyen:

- Cambios en las concepciones acerca de cómo utilizar la tecnología dentro y fuera del aula.
- Definición de los procesos didácticos y la identidad del docente en la actualidad.
- Cambios en los recursos básicos: contenidos, materiales, infraestructuras y el uso de estos recursos.
- Cambios en las prácticas de los profesores y de los alumnos.

La reflexión sobre todo ello debe hacerse por medio del análisis de la disponibilidad tecnológica. Desde la viabilidad económica hasta la viabilidad didáctica, poniendo especial atención en la calidad de los materiales y de los sistemas de enseñanza, junto con las posibilidades que ofrecen dichos sistemas para lograr entornos de aprendizaje efectivos para conseguir la interacción entre profesor y alumno.

3.1.3 Procesos para la generación y publicación de contenido

En la formación a distancia tradicional, la generación de contenido es la fase más importante y decisiva, puesto que a menudo el curso se identifica con los materiales formativos. En cambio, para las nuevas modalidades de formación en línea y los cursos presenciales apoyados en TIC, hay otros componentes que adquieren relevancia cómo es la interacción entre alumnos, el trabajo colaborativo o la posibilidad de interrelacionarse con el tutor en tiempo real, y deben tomarse a consideración al momento de elaborar los contenidos del curso (7).

También ha cambiado el concepto de contenido. Durante mucho tiempo el formato en papel ha sido el dominante, ya fuera acompañado de otros formatos o no, pero la logística de la distribución y el proceso de corrección consumían tiempo y esfuerzo considerables. En la formación asistida por TIC, los materiales se han diversificado y en algunos casos se han integrado en un formato multimedia, pudiendo, ahora, ser distribuir en la propia web o a través de medios de almacenamiento (CD-ROM, memoria flash, tarjeta SD, etc.). Además, gracias a los espacios de comunicación como el chat o foros de debate, un simple texto, junto con la retroalimentación que se pueda obtener en un comentario o debate, se convierte en un material con un valor formativo significativo.

La creación de contenidos empieza en el momento que se presenta el plan de estudios de un curso. Las personas encargadas de la producción de un curso tienen que ser sensibles a las demandas de los alumnos, puesto que el curso requiere establecer los puntos clave (objetivos, características del contenido, tipos de actividades, impartición, evaluación, etc.) que los materiales van a reflejar.

Un modelo general para la generación de contenidos de calidad, puede observarse en la **Figura 1**. El proceso puede ser aplicado tanto en una institución de carácter formativo como en las empresas donde el conocimiento sea un factor de producción.

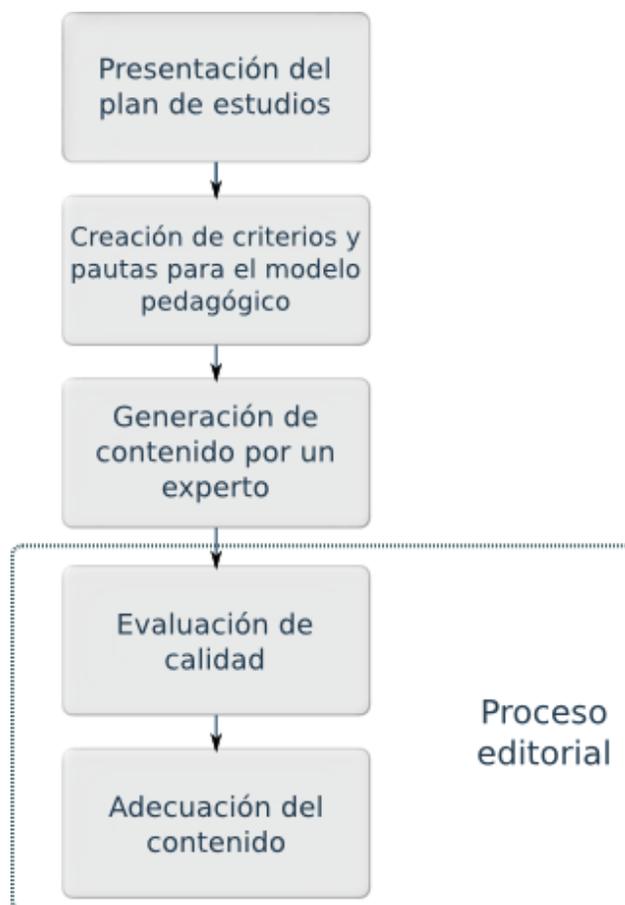


Figura 1 - Proceso para la generación de contenido (8)

Los materiales deben plantearse como una herramienta del propio curso y su creador debe tener siempre en cuenta la plataforma virtual donde se impartirán, siguiendo los estándares recomendados, y no caer en el error de considerar las TIC como simples accesorios de los materiales.

Para que un curso pueda cumplir con sus objetivos, los editores, diseñadores y productores de materiales deben procurar que sea atractivo para el alumno, que facilite el estudio y el aprendizaje, empleando los recursos que tienen a su alcance de tal manera que resulte beneficioso para el

proceso de creación. Así pues, los contenidos de un curso irán en el formato más idóneo en cada caso (texto, audio, video, simulación, web, multimedia).

Los expertos generalmente redactarán textos ya sea en formato papel o en formato digital; difícilmente se les podrá pedir que programen, que graben vídeo, que elaboren sus propias imágenes o que diseñen un sitio web; para ello tendrán que intervenir otros profesionales, como ingenieros y diseñadores.

El primer paso consiste en garantizar que los contenidos redactados correspondan con la orientación académica y el plan de estudios.

Hay cursos que están basados en contenidos ya existentes, como libros, manuales y vídeos. En general, estos materiales se acompañan de una guía didáctica que orienta el trabajo del alumnado. Esta práctica es la predominante, pero presenta diversos problemas (9). Por ejemplo:

- Los materiales ya editados podrían no tener la extensión necesaria para cubrir lo que se ha establecido en el temario.
- Es muy probable que existan algunos vacíos, los cuales el profesor se verá obligado a cubrir.
- Si el material se toma de diversas fuentes, existe el riesgo de que los recursos recabados se solapen, al abarcar los mismos temas.
- En algunas situaciones, exigen un esfuerzo más grande por parte del alumnado o el profesor, para adecuarse al material pero sin dejar a un lado los objetivos del curso..

Es preferible establecer modelos de creación y edición propios de la institución formativa que se adecuen a sus necesidades y creen un modo distintivo de generar contenido, y por tanto, conocimiento. Tanto los modelos como el contenido deben estar en constante evolución. Para ello, la institución puede implementar diversas prácticas y elementos (8):

- Un estilo y visión propios de la edición de los contenidos didácticos en línea.
- Un modelo de calidad propio y una determinación de qué es aquello que aporta valor añadido en este proceso.
- Criterios y pautas para establecer dicho modelo pedagógico.
- Criterios de calidad que deben cumplir los programas formativos y, en general, los contenidos que van a formar parte de los mismos.
- Documentos que definan el trabajo, objetivos, características e instrucciones de todo el proceso, para que terceros lo puedan llevar a cabo con el máximo grado de precisión.
- Estrategias y herramientas para que los procesos sean más eficientes.

La edición (formalmente llamado proceso editorial) es el siguiente paso dentro del proceso de generación de contenidos. La edición de contenidos debe garantizar que éstos se ajusten al diseño del curso planteado, para ello, se necesita reconocer cuáles contenidos requerirán de intervención para dotarlos de un diseño adecuado, y el grado en que serán modificados; debe existir un margen bien definido acerca de las manipulaciones que se permitirán sobre dichos contenidos.

Para que un experto pueda plasmar su conocimiento de la forma más didáctica posible, con un formato, estructura, objetivo y estilo determinado se le proporcionan las pautas y criterios del modelo pedagógico. Pero normalmente, este autor se queda en un nivel didáctico concreto y no será requisito obligatorio que dé respuesta a la aplicación de todas las técnicas didácticas que se requieren en la formación asistida por TIC (10).

De este modo, puede suceder que el contenido original, aun siendo un trabajo formidable en el que se cumplan los requisitos establecidos, requiera de ajustes para que dicho contenido se adecue a la secuencia didáctica establecida y al medio en el que se representará, de modo tal que genere aprendizaje y sea coherente con los contenidos existentes.

No obstante, es importante aclarar que durante el proceso de edición no se trata de reescribir la totalidad de un material, de lo que se trata es de adecuarlo y desarrollar toda su potencialidad.

La adecuación de un contenido requiere de un previo análisis del mismo, en donde se reconocerán las modificaciones pertinentes y se dotará de calidad y carácter pedagógico al documento original. Para la evaluación de la calidad se pueden usar diferentes mecanismos e indicadores, el más usado y aceptado en la actualidad es la *revisión por expertos* (conocido como *peer review*^[1], por su designación en inglés) (11), el cual genera un ambiente de calidad y neutralidad. También permite mejorar la presentación del trabajo, eliminar errores, evitar omisiones e incrementar el potencial didáctico del contenido.

3.1.4 Revisión por expertos

El desarrollo de esta tesis está estrechamente relacionado con el concepto de revisión por expertos, puesto que es el método que será implementado durante el proceso de evaluación de calidad. Por tanto, es necesario hacer hincapié en esta modalidad de revisión.

Cuando la comunidad académica confía en la calidad de las publicaciones o contenidos académicos lo hace, en parte, por el proceso de evaluación que está implicado.

La función de los expertos, que dentro del proceso son llamados revisores, consiste en evaluar el trabajo presentado y emitir un dictamen de entre cuatro posibles opciones (12):

¹ Es común encontrar este concepto traducido directamente del inglés como “revisión por pares”, aunque dicha traducción no expresa que la evaluación será realizada por una persona considerada como profesional, dentro del campo que compete al material a revisar, no solo por un par o coetáneo.

1. Aceptar el trabajo sin ninguna modificación.
2. Aceptar el trabajo con cambios menores (cambios que harán los editores o los mismos revisores, con consentimiento del autor).
3. Devolver el trabajo junto con una lista de modificaciones sugeridas o errores que el revisor haya encontrado. Se espera que el autor haga caso de las correcciones y entregue un trabajo corregido.
4. Rechazar el trabajo.

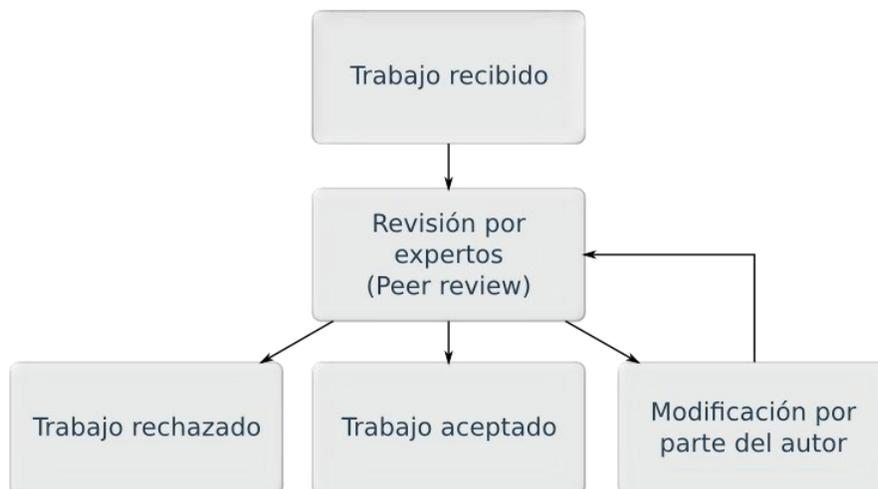


Figura 2 - Proceso de revisión

Los revisores, cuando se trata de instituciones formativas, generalmente no reciben ninguna remuneración económica por su labor. Pero debe tomarse en cuenta que el ser elegido como experto para evaluar un trabajo supone un reconocimiento y confiere cierto prestigio. Por otra parte, cuando se revisa un trabajo original, se tiene un acceso privilegiado a información. Además, muchos consideran que la revisión de los trabajos es uno de los deberes básicos de la comunidad formativa, siempre con el fin de contribuir a aumentar el conocimiento que se genera y dotarlo de calidad.

El sistema de revisión por expertos consiste en que dos o más revisores lean y analicen los artículos para determinar tanto la validez de las ideas y los resultados, como su impacto en el área que abordan (11).

Los contenidos son recibidos por la institución formativa, quien los envía a los revisores sin el nombre del autor ni su filiación, para evitar distorsiones en la evaluación. La institución recibe el informe de los revisores y cuando informan al autor el dictamen, no detallan, tampoco, el nombre

de los revisores ^[2]. Aunque, en algunos casos la revisión es abierta, es decir, autores y revisores conocen su identidad, pero no es una modalidad común.

En aquellos casos en que surja una gran disparidad en los informes de los revisores, la institución suele recurrir a la opinión de un revisor externo.

La revisión de trabajos por expertos ha adquirido gran importancia hasta convertirse en una herramienta clave del progreso de muchas áreas, como lo son la ciencia, pedagogía y periodismo. Ha sido reconocida como un buen método para imponer uniformidad en las publicaciones y garantizar la calidad de éstas. Pero aun así se debe recordar que basarse en el proceso de revisión por expertos no garantiza de forma absoluta la calidad de los contenidos.

² Este proceso en donde ninguna de las partes (autores y revisores) conoce el nombre o filiación de aquellos implicados en la revisión, se conoce como “revisión a doble ciego”.

3.2 Procesos de gestión de información

La creciente importancia de la información como un nuevo factor de producción hace que el desarrollo de tecnologías, metodologías y estrategias para su creación, medición y difusión se convierta en una prioridad. Sin embargo, también podemos considerar que ha sido precisamente el desarrollo de esas tecnologías y metodologías el que ha convertido la información, y por tanto el conocimiento, en un elemento indispensable para el desarrollo económico, social y cultural.

La gestión de información ha surgido en las empresas y en las instituciones formativas a comienzos de los 90's, debido a tres razones (13):

1. Rápida obsolescencia de las bases de conocimiento.
2. Urgencia por valorar intangibles.
3. Necesidad de integrar soluciones de tecnologías de la información.

Existen múltiples enfoques relacionados con la gestión de información y, por tanto, múltiples definiciones. Sin embargo es necesario definir lo que, en lo pertinente a esta tesis, se entiende por gestión de información. Para ello se ofrecen dos definiciones que, en conjunto, ofrecen un panorama general de lo que implica la gestión de información:

«La gestión de información se puede definir como un proceso sistemático de búsqueda, selección, organización y difusión de información, cuyo objeto es aportar a los profesionales de la compañía los conocimientos necesarios para desarrollar eficazmente su labor» (14)

«La gestión de la información es el proceso de administrar continuamente información de todo tipo para satisfacer necesidades presentes y futuras para identificar y explotar recursos de conocimiento tanto existentes como adquiridos y para desarrollar nuevas oportunidades» (15)

Cabe destacar que las definiciones expuestas están centradas en la gestión de información dentro de una compañía u organización lucrativa. No obstante, tales conceptos pueden ajustarse sin problemas al contexto de las instituciones formativas. Aunado a ello, es notable cómo ambas definiciones convergen en un detalle significativo, que es considerar a la gestión de información como un proceso.

Efectivamente, la gestión de información es todo un proceso cíclico conformado por diversas fases. La información pasa por cada una de estas fases durante el proceso de gestión, y cada fase es por sí misma un proceso. En general, se reconocen 7 fases dentro del proceso de gestión (16) (15) y cada una será explicada posteriormente:

1. Identificación
2. Creación
3. Almacenamiento
4. Estructuración
5. Distribución
6. Mantenimiento
7. Contabilización



Figura 3 - Ciclo de gestión de información (16)

Identificación de información

El primer paso dentro del ciclo es identificar la información que reside en la institución, así como sus características. Este proceso constituye el núcleo del ciclo, es el elemento necesario para poner en marcha la gestión de la información, ya que cualquiera de las fases girará siempre en torno a la información identificada.

Ahora bien, el proceso de identificación de la información también implica reconocer los procesos que la institución requiere para hacer uso de la misma, con el fin de diseñar las técnicas de formación, desarrollo y retribución del recurso humano (autores, editores, diseñadores, etc.) (15).

Al inicio se debe crear una lista preliminar con toda la información reconocida que se considere relevante para la institución, dependiendo del enfoque que se tenga, o mejor dicho, la necesidad que exista por gestionar la información.

Una vez obtenido el listado, la información que tenga contenido similar se agrupará, con el fin de evitar duplicidades o redundancias.

Finalmente, una vez agrupada la información, se realizará una definición preliminar de cada uno de los grupos detectados. Para ello se podrá contar con expertos en cada una de las materias.

Los expertos validarán el listado con el fin de asegurar que no se ha dejado fuera ningún dato o proceso relevante, y también, que no se haya incluido alguno que no debería estar. Por ello, este grupo de expertos debe pertenecer a la institución ^[3] y un tener claro conocimiento de la visión global de la institución.

Creación de información

La creación constituye el verdadero inicio de la gestión de información. La identificación debe considerarse como un paso previo, pero en diversos contextos (sobre todo empresariales) su relevancia es tal que el mismo proceso de gestión no podría dar inicio sin la identificación de la información existente. Dentro de esta etapa existen todas las actividades y procesos encaminados a crear nueva información, tanto individual como colectiva. (16)

Para una institución formativa, los procesos de generación de información (y de conocimiento) son actividades que habitualmente se realizan en las actividades de formación tradicional. Aunque existe una diferencia significativa, ya que la información se debe poner a disposición de todos los implicados en el proceso de gestión.

Si bien el proceso de creación en una institución formativa es inherente a la misma; para una empresa resulta un problema, puesto que, en el ambiente empresarial, la creación de información no es un proceso que se planifique y controle, tan solo se encuentra implícito dentro del proceso de innovación. Por ello, una estrategia que se suele seguir en las organizaciones es el de prueba y error. Es decir, mediante un proceso iterativo de pruebas, se extrae información y conocimientos parciales, los cuales se estructuran para enfocar el proceso de creación en el camino correcto (17).

Se recomienda motivar al personal a buscar la información que consideren necesaria para el desempeño de sus actividades y el crecimiento de la institución. Para ello es necesario capacitarlos, de modo que no sólo sepan cómo generar información y plasmar su conocimiento, sino que sean capaces de encontrar más información sobre los aspectos de su interés y aprender a buscar nuevas fuentes fidedignas de información.

La capacidad individual de generar y buscar información es crítica para conseguir una adecuada gestión de la misma.

³ Esta afirmación pudiera sonar obvia, pero es un error común el depender de personal externo a la organización, sobre todo si se trata de solicitar la opinión de un experto.

Almacenamiento de información

Una vez creada (o identificada) la información, el siguiente paso dentro del ciclo es almacenarla. Al hablar de almacenar la información, antes, se debe hacer referencia a dos procesos:

1. Combinación.
2. Conversión.

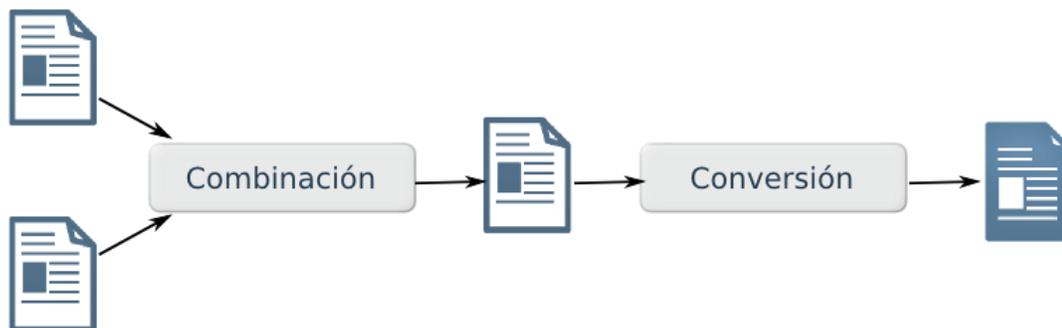


Figura 4 - Procesos previos al almacenamiento

La combinación de información, puede verse como una adecuación, y posible síntesis, de la información existente. Se da lugar en aquellos documentos que contengan información similar y se desee evitar redundancias. También, se da lugar cuando se descubre una fuerte dependencia entre dos documentos, en donde sea necesario hacer referencia a uno para poder comprender el contenido del otro; la combinación facilitaría la comprensión de todo el conjunto de información comprendido entre ambos documentos.

La conversión de información es el proceso en donde se traducen los documentos existentes de manera que puedan ser entendidos por todos los demás implicados en el proceso de gestión. Asimismo, es necesario transformar el medio (físico o lógico) que contiene la información, a fin de poder almacenarla adecuadamente. Dicha transformación debe realizarse en un formato que permita que la información pueda estructurarse, transmitirse y utilizarse de manera sencilla y, sobre todo, sin tener que recurrir a la persona o personas que crearon la información (17).

Para poder almacenar la información, se dispone de una gran variedad de medios para codificarla: en papel, en archivos de audio, en archivos multimedia, etc. Sin embargo, el proceso de almacenamiento de información no consiste tan solo en la codificación del mismo, sino que será necesaria su interpretación, para darle un sentido coherente antes de hacerla pública, aun cuando solo sea utilizada internamente por la institución u organización. Debe considerarse que la información se debe almacenar en formatos que sean fácilmente accesibles por parte de toda la organización. Es en este punto donde las TIC juegan un papel decisivo, facilitando el proceso de almacenamiento y distribución.

Las TIC ofrecen claras ventajas sobre otros medios más tradicionales. Tecnologías como hipertexto o multimedia permiten crear y almacenar conocimientos a gran escala y en un menor

espacio. Además, el acceso al conocimiento se puede personalizar y diseñar a la medida, dependiendo del nivel de conocimientos y de las necesidades de cada persona. Los miembros de una institución pueden acceder a la información necesaria de manera independiente y en el momento que lo requieran. Inclusive, el empleo de TIC, permite la normalización de los contenidos formativos y facilitan su actualización.

Una de las soluciones provistas por las TIC (en realidad un conjunto de herramientas) es el Sistema de Gestión de Contenidos (Content Management Systems, CMS) que, por sus características funcionales propias, llega incluso a considerarse como el estándar tecnológico *de facto* para la implementación de los procesos de gestión de información (14).

Estructuración de información

En la práctica, se considera que almacenar y estructurar son una sola fase dentro del ciclo, puesto que de nada sirve tener información almacenada si posteriormente no podrá ser encontrada. Por tanto, se trata de dos funciones que deben realizarse en paralelo.

Al momento de estructurar la información se deben establecer (o incluso deducir) cuales son los criterios o pautas para clasificar la información. No existe realmente una guía o metodología a seguir para formular estructuras, depende del tipo de información y enfoque de la institución, aunque una estructuración funcional debe tomar en cuenta los siguientes aspectos:

- La estructura debe ser coherente con el tipo de información a clasificar.
- Debe ser sencillo realizar búsquedas de información.
- Aun sin un método de búsqueda, debe ser posible encontrar la información deseada por medio de la clasificación que se proponga.
- La estructuración debe atender a las necesidades de aquellos que requieran la información.

Es conveniente realizar la estructuración de forma matricial. Esto quiere decir que en un eje iría la clasificación interna que se haya decidido y en el otro la información existente, como se ejemplifica en la **Tabla 1**. Sin embargo, no es necesario colocar cada elemento de información en la estructuración matricial, puede agruparse esta información de modo tal que la persona interesada en una categoría en concreto, podría acceder directamente a un grupo de información, siguiendo la clasificación, y desde ahí acceder a las distintas piezas en que se estructura el mismo.

	Categoría 1	Categoría 2	Categoría 3	Categoría 4
Documento A	✓			
Documento B		✓		
Documento C	✓	✓	✓	
Documento D	✓			✓
Documento E			✓	

Tabla 1 - Matriz de Estructuración

Distribución de información

Una vez almacenada y estructurada la información, se debe colocar a disposición de la institución u organización. Es necesario asegurarse que el proceso de gestión y los sistemas de TIC involucrados estén pensados para lidiar con el carácter social del conocimiento.

La información se debe poner en circulación, las personas que accedan a la información podrán incluir sus aportaciones, matizar el contenido existente, modificar información, realizar correcciones, etc., todo ello aportando su propio conocimiento y, por tanto, más información.

Por otro lado, en la distribución de información, hay que tener en cuenta que existen dos variantes:

- La distribución interna: entre los propios miembros de la institución (suele ser la variante más utilizada, sobre todo en empresas).
- La distribución externa: a personas ajenas a la institución.

Mantenimiento de la información

La información ya existente no puede permanecer invariable, también debe evolucionar, para mantenerse al día de los cambios que se producen en el entorno. De lo contrario perderá su valor lo cual volvería a la gestión de información un proceso de poca retribución, es decir, prescindible.

Esto requiere de programas específicos cuya finalidad sea aprovechar la información acumulada, estimular los procesos creativos, de investigación y desarrollo.

El proceso de mantenimiento debe ser asumido por toda la institución. Una vez en marcha todo el ciclo, el mantenimiento debe darse de forma natural a lo largo del mismo, fruto de la creación de nuevo conocimiento y de la distribución y compartición de información.

No solo la información se debe mantener, también los sistemas y herramientas implicadas en el proceso, algunas de las cuales habrán sido diseñadas específicamente para gestionar la información.

El mantenimiento de los sistemas, debe ser responsabilidad directa de las áreas involucradas en el desarrollo o implementación de estos sistemas. Todo cambio en el proceso de gestión debería implicar ciertas modificaciones en el las herramientas de TIC.

Contabilización de información

Los recursos intangibles (como la información o el software) constituyen una parte fundamental del valor de una institución o empresa. Por tanto, es necesario desarrollar indicadores que permitan medir la información para mejorar la gestión, así como para valorar correctamente dichos intangibles.

A este proceso se le denomina contabilización. Aquí es donde se debe valorar la contribución de todas las personas implicadas. La contabilización constituye el cierre del ciclo, ya que este proceso permite medir la efectividad de la gestión de información.

El problema radica en cómo medir lo intangible. En este sentido existen varias iniciativas que se han puesto en marcha con el fin de establecer un sistema que permita realizar, de manera más objetiva, dicha medición, mediante el establecimiento de un sistema de indicadores universales, para conocer el valor de la información.

Para realizar la contabilización de información, se puede hacer uso de modelos enfocados en recursos intangibles y gestión del capital intelectual (18) (19) (20):

- Modelo Navegador, propuesto Leif Edvinsson.
- Cuadro de mando integral, propuesto por Robert S. Kaplan y David P. Norton.
- Monitor de activos intangibles, propuesto por Karl Erik Sveiby.
- Modelo de dirección estratégica por competencias, elaborado por Eduardo Bueno.
- Technology Broker, desarrollado por Anny Brooking.
- MERITUM, creado por la Universidad Autónoma de Madrid y la Universidad de Sevilla, en conjunto con universidades de Francia, Noruega, Suecia, Dinamarca y Finlandia.

Cabe notar que no es forzoso adoptar alguno de los modelos existentes, sean los aquí listados o cualquier otro. Una institución o empresa puede contabilizar su información de la manera que crea conveniente, implantando medidas propias, por muy rudimentarias que pudieran resultar. La importancia en este proceso de contabilización no radica en el modelo a utilizar, sino en asegurarse que existe alguna forma de comprobar el rendimiento de la gestión de información.

3.3 Tecnologías involucradas

3.3.1 Formatos de archivo

Los archivos son un mecanismo de abstracción que permite almacenar información para poder acceder a ella después. Esto debe hacerse de tal modo que el usuario no tenga que enterarse de los detalles de cómo y dónde está almacenada la información.

Un sistema operativo no sabe qué contiene el archivo, únicamente lo ve como una sucesión de bytes que tiene una ubicación y nombre específico. Cualquier significado que tenga el archivo deberán imponerlo los programas a nivel de usuario.

Los archivos pueden estructurarse de varias maneras al ser almacenados. Un formato de archivo es una forma particular en que la información se estructura para su interpretación.

El formato de un archivo incluye un encabezado, datos y, opcionalmente, un marcador de final del archivo. Un encabezado de archivo es una sección de datos al inicio de un archivo que contiene información acerca del mismo: suele ser la fecha de creación, la fecha de su última actualización, su tamaño e información específica del formato mismo.

El resto del archivo depende de los datos que contenga, ya sea de texto, gráficos, audio, etc. Por ejemplo, un archivo de texto puede contener frases y párrafos intercalados con códigos de formato al texto. Un archivo de gráficos puede contener datos sobre el color para cada píxel y una descripción de la paleta de colores. El formato de archivo determina el orden de estos datos.

La **Figura 5** ilustra el diseño para un archivo de mapa de bits de Windows (BMP) y lo compara con el diseño de un archivo en formato Graphics Interchange Format (GIF). Ambos formatos permiten almacenar imágenes, para el caso del formato GIF, es posible almacenar más de una imagen dentro de un mismo archivo, lo cual es utilizado frecuentemente para crear animaciones sencillas.

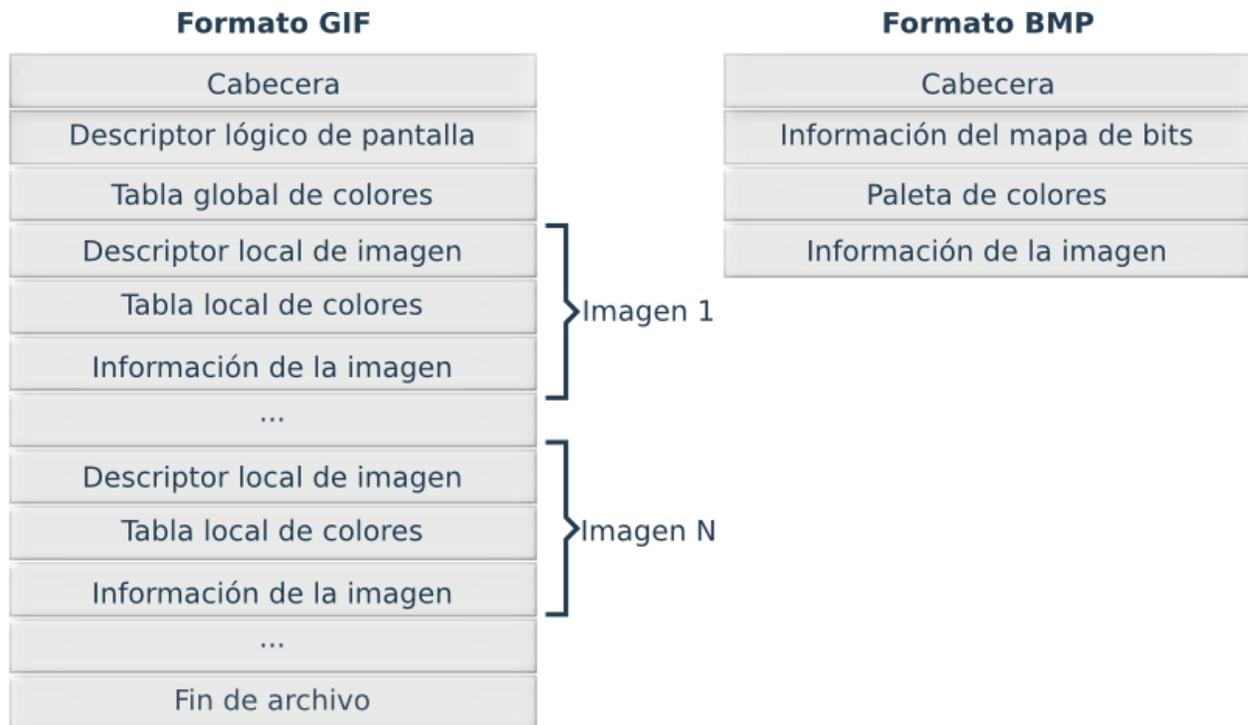


Figura 5 - Formatos de imagen

Algunos sistemas operativos funcionan de tal modo que protegen a los usuarios de las complejidades de los formatos de archivo. Por ejemplo, Windows emplea una lista de asociación de archivos para relacionar una extensión de archivo con su software de aplicaciones correspondiente. Esta función le permite abrir un archivo sin abrir primero una aplicación.

Los sistemas operativos no proporcionan ningún mecanismo formal para identificar el formato de un archivo. Sin embargo, existen dos elementos convencionalmente usados para poder informar a las aplicaciones cuál es el formato de un archivo:

- Extensión del archivo
- Número mágico

Extensión de archivo

En algunos sistemas operativos, como en MS-DOS, Windows y Unix, se utiliza parte del nombre para definir el formato de un archivo, donde la extensión del archivo son los últimos caracteres del nombre, después del punto. Por ejemplo, en el archivo *reporte.txt*, la extensión es *txt*, e indica que el archivo es un documento de texto.

La extensión, en algunos sistemas operativos como Windows, indica qué programa debería ser usado para abrir el archivo. También es útil para que el usuario conozca qué información puede tener un archivo sólo con mirar su nombre.

Ésta es la base del sistema MIME ^[4]. Tiene el problema de que cualquier usuario o software puede modificar fácilmente la descripción del archivo. En Mac OS, en las primeras versiones, se almacenaban el formato y el creador como metadatos del archivo. Sin embargo, estos eran libremente modificables por un programador. La estrategia que han asumido la mayoría de los formatos estándares en Internet es utilizar un encabezado en el archivo, el cual se ajusta a ciertas convenciones.

Número mágico

El número mágico, también conocido como firma del archivo, se encuentra ubicado en los primeros bytes del archivo (en el encabezado), indicando su formato. Se llama número mágico porque su propósito y significado no es aparente si no se tiene un conocimiento adicional. Se puede usar un visor hexadecimal para acceder a este número de forma manual que, generalmente, para los formatos de archivo más comunes, representa el nombre del formato de archivo.

No todos los formatos cuentan con un número mágico; los archivos de texto plano, como los HTML (hypertext markup language), los XHTML (extensible HTML), los XML (extensible markup language) y los archivos de código fuente no están identificados con una firma de archivo.

La **Tabla 2** muestra algunos ejemplos de números mágicos, tanto su representación hexadecimal como los caracteres ASCII correspondientes.

⁴ Multipurpose Internet Mail Extensions, es un conjunto de convenciones para el intercambio de archivos a través de Internet.

Formato	Extensión	Número mágico hexadecimal *	Representación ASCII **
Mapa de bits de Windows	BMP	42 4D	BM
Imagen de disco compacto ISO-9660	ISO	43 44 30 30 31	CD001
Waveform Audio Format	WAV	52 49 46 46 xx xx xx xx 57 41 56 45 66 6D 74 20	RIFF....WAVEfmt
Archivo comprimido WinRAR	RAR	52 61 72 21 1A 07 00	Rar!...
Archivo de enlace de Windows	LNK	4C 00 00 00 01 14 02 00	L.....
Music instrument digital interface	MID, MIDI	4D 54 68 64	MThd
Archivo de video Flash	SWF	46 4C 56	FLV

* Una equis 'x' simboliza que el número puede variar.

** Un punto '.' Simboliza que no existe representación ASCII para el número hexadecimal.

Tabla 2 - Ejemplos de números mágicos

3.3.2 Acervos digitales

Los acervos digitales se han convertido en herramientas esenciales para el desarrollo académico. Las definiciones que se han dado de acervo digital han sido numerosas. Entre ellas, podemos destacar las siguientes:

«Una colección digital que agrupa y preserva la producción intelectual de una o varias universidades» (21). Esta definición establece un campo de aplicación muy amplio, que engloba el conjunto de la producción intelectual de la institución, no solo los resultados de investigación y objetos de aprendizaje sino también cualquier otro producto de carácter cultural producido por la institución.

«Un conjunto de servicios que una universidad ofrece a los miembros de su comunidad para la gestión y diseminación de los materiales digitales creados por la institución y sus miembros» (22). Es importante notar la consideración de servicio que se da al acervo, como algo dinámico al servicio de la comunidad, más allá de un depósito pasivo de documentos digitales.

Una definición más abierta considera que un acervo digital debe cumplir al menos las siguientes características (23):

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

De estas definiciones se puede concluir que un acervo digital es un conjunto de servicios para para recopilar, administrar, difundir y preservar la producción documental digital generada en la institución, cualquiera que sea su tipo (investigación, docente, cultural, etc.).

Un acervo puede formar parte de un sistema mayor, interconectándose con otros acervos a nivel regional, nacional o global; indizados de una manera estándar para garantizar un aumento de la visibilidad e impacto de la misma. También pueden proveer las bases para nuevos modelos de edición académica y pueden vincularse a otros servicios relacionados, como enseñanza a distancia (e-learning) o publicación electrónica de revistas (22).

Repositorios institucionales

La diversidad de formatos y tipos de documentos desarrollados, la multitud de funciones demandadas por las instituciones y la gran posibilidad de aplicaciones ofrecidas por las TIC ha generado cierta ambigüedad en el uso de los términos que definen al sistema encargado de generar, proteger y difundir la producción de una institución.

Cada vez es más común escuchar el término *repositorio institucional (RI)* al hablar de cualquier tipo de colección de documentos digitales, pero antes de realizar una comparación o aclaración con los acervos digitales, es necesario recordar que los RI surgen como una respuesta de las instituciones, en especial las académicas, hacía la política de las revistas científicas tradicionales, tendente hacía la subida constante de precios, y la necesidad de las instituciones de conservar, preservar y poner a disposición de su comunidad académica e investigadora su patrimonio intelectual (21). En respuesta a este panorama de la publicación científica comercial surgen los dos ejes fundamentales en donde debe basarse todo proyecto de repositorio institucional:

- La difusión de la investigación.
- El acceso abierto a los documentos.

El formato de presentación del documento no debe ser determinante. Por ejemplo: tesis, artículos de revistas, fotos, mapas, documentos económicos o financieros, estadísticas, etc., pero también ideas, propuestas, hipótesis, experimentos, datos e informe de resultados entrarían dentro de un RI siempre que sean el resultado de una investigación institucional. Estos serían los factores clave identificadores de un repositorio institucional, a diferencia de cualquier otro tipo

de sistema de gestión de contenidos cuyas características pueden ser totalmente diferentes y, por supuesto, estar más orientado hacia la enseñanza-aprendizaje, como es el caso de los acervos digitales.

En términos generales se puede definir a un repositorio institucional como «*colecciones digitales de captura y preservación de la producción intelectual de una universidad o comunidad de múltiples instituciones de colegios y universidades*» (24).

En esta definición se habla de *producción intelectual* sin especificar si se trata de resultados de investigación o también de materiales destinados a la docencia.

En este sentido, se pueden analizar las dificultades presentes en la aplicación de la misma filosofía de trabajo de un RI a un material didáctico:

- Sí se admiten los materiales didácticos a un IR, también debería hacerse con otros tipos de documentos orientados a la docencia como los libros de texto, apuntes, programas curriculares, etc., sobre los cuáles es difícil mantener un control por parte de un RI.
- Los materiales didácticos varían enormemente en formatos, requerimientos de uso de metadatos, tamaño, etc.
- Los materiales didácticos aún no están muy difundidos y, menos aún conservados o accesibles en repositorios institucionales. Esto supone un inconveniente a la hora de desarrollar un repositorio institucional de materiales didácticos a la medida del contenido ya existente.
- Las plataformas de software desarrolladas para los materiales didácticos aún no están normalizadas debido a las características de este tipo de material, como lo puedan estar para los resultados de investigación.

Un RI y un acervo digital están definidos no tanto por el formato de los documentos constituyentes de su contenido sino por su finalidad: «*Los repositorios institucionales son, tal vez, particularmente aplicables en el contexto de las publicaciones de investigación, puesto que emanan de las instituciones, y con la tecnología adecuada pueden convertirse en servicios*» (25)

Por otro lado, muchos autores e instituciones defienden la enseñanza-aprendizaje como una de las funciones clave de los acervos digitales. El uso docente de los acervos digitales va más allá de la difusión de los resultados de investigación de los IR. Se trata de un sistema de gestión de información, cuyos objetivos son docentes y cuyo último destino no tiene por qué ser la publicación y la difusión de los resultados generados por la institución.

3.4 Bases de datos

La información clara, oportuna y útil requiere datos precisos. Tales datos deben generarse apropiadamente y deben guardarse en un formato que sea fácil de acceder y procesar. Y, al igual que cualquier recurso básico, el ambiente de los datos debe manejarse con cuidado. Dado el rol crucial que desempeñan los datos, su administración es una actividad primordial en cualquier institución.

En general, la administración de datos eficiente requiere el uso de una base de datos de computadora. Existen diversas definiciones para describir una base de datos, aunque muchos autores prefieren definir lo que es un sistema de gestión de base de datos, y dejar la definición de la propia base de datos abierta a interpretación o denotarla simplemente como un depósito de datos organizado y de fácil acceso.

Aun así, es preferible hacer la correcta distinción entre la base de datos y su sistema gestor, para ello es necesario enfocarse en dos definiciones significativas:

«Una base de datos es un conjunto estructurado de datos que representan entidades y sus interrelaciones. Dicha representación será única e integrada, permitiendo usos diversos y simultáneos.» (26)

Esta definición nos permite saber que los datos almacenados en una base de datos no solo son cifras y texto organizado, sino que es la representación de un objeto o acontecimiento del mundo real del cual se desea guardar información. Toda la información que se almacene dentro de una sola base de datos debe tener algún tipo de relación entre sí, es decir debe ser coherente bajo el contexto donde se sitúa.

«Una Base de datos es un conjunto de datos almacenados sin redundancias innecesarias en un soporte informático accesible simultáneamente por distintos usuarios y aplicaciones. Los datos deben de estar estructurados y almacenados de forma totalmente independiente de las aplicaciones que la utilizan.» (27)

La segunda definición expone que una base de datos tiene un soporte informático que separa los datos de las aplicaciones, de este modo no dependen de la aplicación del usuario final, y por ello se pueden almacenar datos cuya estructura pueda estar definida conforme a nuestras necesidades, ya que será el sistema de gestión el que se encargue de conectar la base de datos con las aplicaciones de los usuarios finales.

Una base de datos, en general, aloja dos conjuntos de objetos:

- Datos: es la representación de una entidad o suceso, ya sea como número, texto, símbolo, etc. Son el componente fundamental de una base de datos. Un dato por sí mismo no aporta conocimiento, para ello debe ser interpretado y procesado.
- Metadatos: la acepción más ampliamente utilizada indica que un metadato es el dato sobre un dato, es decir, la información mínima necesaria para identificar un recurso. En el capítulo 3.5 se explicará a detalle esta definición.

El sistema de base de datos se compone de cinco grupos principales, mostrados en la **Figura 6**: hardware, software, personas, procedimientos y datos.

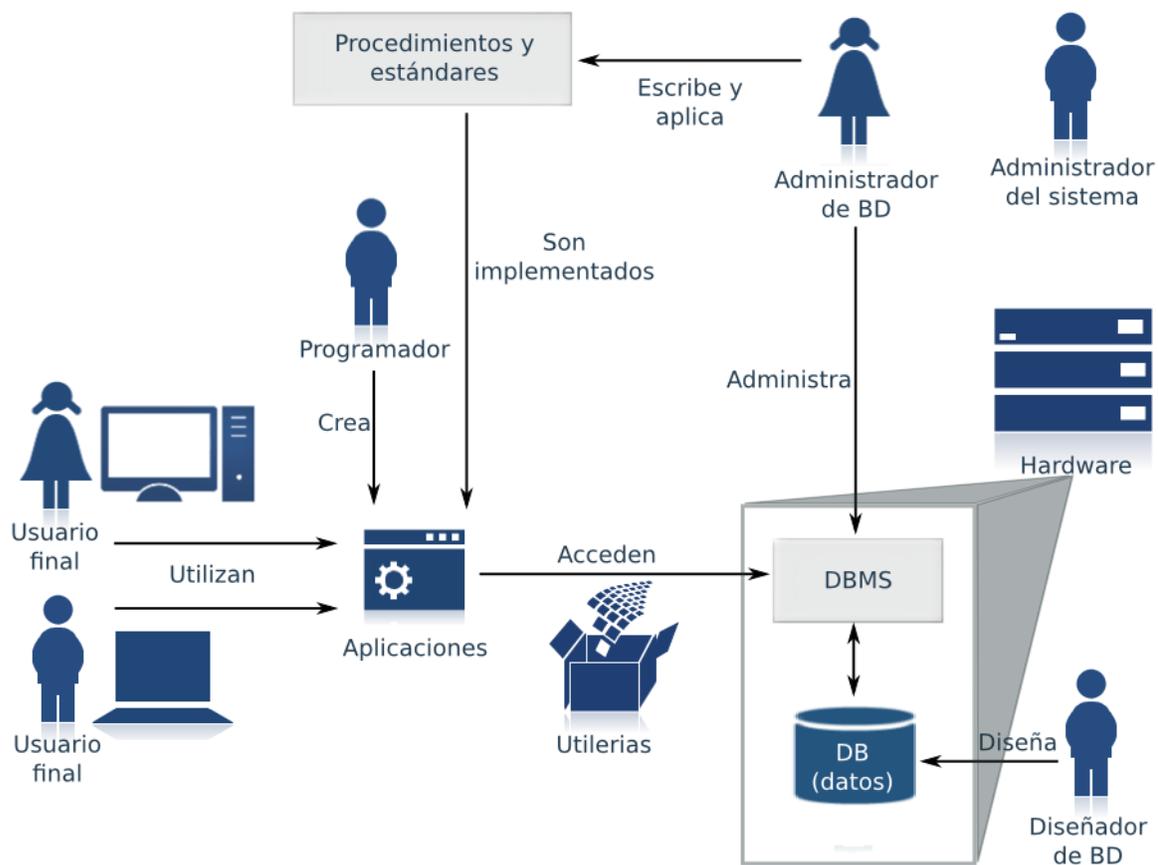


Figura 6 - Sistema de base de datos

Hardware

Se refiere a todos los dispositivos físicos del sistema. Los componentes de hardware del sistema constan de:

- La memoria secundaria de almacenamiento que se emplea para contener los datos almacenados, junto con los dispositivos asociados de E/S y los controladores de dispositivos.
- Los procesadores de hardware y la memoria principal asociada usados para apoyar la ejecución del software del sistema de base de datos.

El hardware principal y más fácil de identificar del sistema de base de datos es la computadora, ya que contiene todos los elementos anteriormente mencionados y puede ser desde una computadora personal hasta un mainframe.

Software

Se refiere al conjunto de programas utilizados por las computadoras dentro del sistema de base de datos. Aunque el software más fácil de identificar es el *sistema de gestión de bases de datos* (DBMS, por sus siglas en inglés), se requieren tres tipos de software para hacer que la base de datos funcione a plenitud:

1. Sistema operativo.
2. DBMS.
3. Programas de aplicación y utilerías.

El software del sistema operativo maneja todos los componentes de hardware y hace posible que todos los demás programas funcionen en la computadora. Algunos ejemplos de software de sistema operativo incluyen sistemas Windows, Linux, MacOS o Z/OS para los mainframes.

El software de DBMS maneja la base de datos y las interacciones entre los datos, las aplicaciones y usuarios. Algunos ejemplos son Access, Oracle, PostgreSQL y SQL Server de Microsoft. Dado que el manejador de bases de datos es uno de los elementos más importantes del sistema, en el capítulo **3.4.1** se detallará su definición y funciones.

Los programas de aplicación y las utilerías se emplean para tener acceso y manipular los datos contenidos en la BD y para manejar el ambiente de la computadora donde se lleva a cabo el acceso y la manipulación de los datos. Los programas de aplicación son elementos de software diseñados para proveer un acceso predefinido a la base de datos, toda la comunicación entre el programa y el DBMS es transparente para los usuarios. Las utilerías son las herramientas de desarrollo de aplicaciones que los programadores utilizan para manejar los componentes de la computadora que intervienen en el DBMS.

Personas

Este componente hace referencia a todos los usuarios del sistema de base de datos. Con base en la función principal de puesto de trabajo, se pueden identificar cinco tipos de usuarios en un sistema de base de datos (28), los cuales se pueden observar en la **Figura 6**:

1. Administradores de sistemas.
2. Administradores de base de datos.
3. Diseñadores de base de datos.
4. Programadores de sistemas.
5. Usuarios finales.

Los administradores de sistemas supervisan las operaciones generales del sistema de base de datos. Los administradores de base de datos (DBA, por sus siglas en inglés), administran el uso del DBMS y garantizan que la base de datos esté funcionando apropiadamente.

Los diseñadores de base de datos proyectan la estructura del sistema. Si el diseño de la base de datos es deficiente, incluso los mejores programadores de aplicaciones y los DBA más dedicados no serían capaces de producir un ambiente de base de datos útil.

Los programadores diseñan y ejecutan los programas de aplicación en algún lenguaje de programación como C#, Java, Ruby, etc. Estos programas acceden a la base de datos emitiendo la solicitud apropiada al DBMS (por lo regular una instrucción SQL). Los programas en sí pueden ser aplicaciones convencionales o pueden ser aplicaciones en línea, cuyo propósito es al usuario final el acceso a la base de datos desde una computadora fuera del área donde se aloja la base de datos.

Los usuarios finales son las personas que utilizan los programas de aplicación para realizar las operaciones diarias de la organización. Un usuario final puede acceder a la base de datos a través de las aplicaciones en línea, mencionadas anteriormente, o bien puede usar una interfaz proporcionada por el DBMS como parte integral del software del sistema de base de datos.

3.4.1 Sistema de gestión de bases de datos (DBMS)

En cualquier tipo de institución se dedican considerables recursos a la recolección, clasificación, procesamiento e intercambio de datos basados en procedimientos bien establecidos para alcanzar objetivos específicos. La gestión o manipulación de estos datos es llevada a cabo por los sistemas de gestión de bases de datos (DBMS).

Los DBMS son sistemas de software, centralizados o distribuidos, que ofrecen facilidades para la definición de bases de datos, la selección de las estructuras de datos (para el almacenamiento y búsqueda de los mismos), tanto interactivamente como mediante un lenguaje de programación.

Un DBMS realiza varias funciones importantes que garantizan la integridad y la consistencia de los datos de una base de datos. La mayoría de estas funciones son transparentes para los usuarios finales, y casi todas pueden realizarse sólo mediante un DBMS. Estas funciones incluyen la administración de un diccionario de datos (metadatos), la administración del almacenamiento de

datos, transformación y presentación de los datos, administración de la seguridad, control de acceso a usuarios múltiples, administración de respaldos y recuperación, administración de la integridad de los datos, lenguajes de acceso a base de datos e interfaces de programación de aplicaciones e interfaces de comunicación con base de datos.

Administración del diccionario de datos (metadatos): el DBMS necesita que las definiciones de los elementos de datos y sus relaciones se guarden en un diccionario de datos. El DBMS utiliza el diccionario de datos para buscar las estructuras y relaciones del componente de datos requeridos. Además, cualquier cambio que se realice en la estructura de una base de datos, automáticamente queda registrado en el diccionario de datos, por lo que el usuario no tiene que modificar todos los programas que tienen acceso a la estructura modificada. En resumen, el DBMS proporciona abstracción de los datos y elimina la dependencia estructural y de los datos del sistema.

Administración del almacenamiento de datos: el DBMS crea las estructuras necesarias para el almacenamiento de datos, con lo que se libera al usuario de la tarea de definir y programar las características físicas de los datos. Algunos DBMS permiten almacenar no sólo datos, sino también, formularios de ingreso o definiciones de filtración de datos relacionados, definiciones de reportes, reglas para validación de datos, código de procedimientos, estructuras para manejar formatos de video e imágenes, etc.

Transformación y presentación de datos: el DBMS transforma los datos que se introducen de acuerdo con la estructura necesaria para guardarlos. Al mantener la independencia de los datos, el DBMS transforma las solicitudes lógicas en comandos que localizan físicamente y recuperan los datos solicitados. Es decir, da formato a los datos físicamente recuperados para que se ajusten a las expectativas lógicas del usuario. En otras palabras, proporciona programas de aplicación con independencia de software y abstracción de datos.

Administración de la seguridad: el DBMS crea un sistema de seguridad que resguardan la privacidad de los datos dentro de la base de datos. Las reglas de seguridad determinan qué usuarios pueden tener acceso a la base de datos, a qué datos puede tener acceso cada usuario y qué operaciones (leer, agregar, eliminar o modificar) puede realizar éste.

Control de acceso de usuarios múltiples: el DBMS crea las estructuras que permiten el acceso de usuarios múltiples a los datos. Para asegurar la consistencia y la seguridad de los datos, el DBMS utiliza algoritmos complejos para permitir el acceso de usuarios múltiples a la bases de datos sin comprometer su integridad.

Administración de tareas de respaldo y recuperación: el DBMS proporciona procedimientos de respaldo y recuperación para garantizar la seguridad e integridad. Los DBMS actuales proporcionan utilerías especiales que permiten que el DBA realice procedimientos de respaldo y recuperación rutinarios y especiales.

Administración de la integridad de los datos: el DBMS promueve y hace cumplir las reglas necesarias para eliminar los problemas de integridad de los datos, con lo que se reduce al mínimo la redundancia de los datos y se incrementa al máximo la consistencia de éstos. Las relaciones entre los datos guardados en el diccionario se utilizan para hacer que se mantenga la integridad de éstos.

Lenguajes de acceso a base de datos e Interfaces de programación de aplicaciones: el DBMS permite el acceso a los datos mediante un lenguaje de consulta. Un lenguaje de consulta es aquel que no tiene procedimientos, es decir, el que permite al usuario especificar únicamente que debe hacer y no como debe hacerlo. El lenguaje de consulta de DBMS contiene dos componentes: un lenguaje de definición de datos (DDL, por sus siglas en inglés) y un lenguaje de manipulación (DML, por sus siglas en inglés). El DDL define las estructuras donde se alojan los datos y el DML permite que los usuarios extraigan los datos de la base de datos. El DBMS también permite que los programadores accedan a los datos mediante lenguajes de programación.

Interfaces de comunicación de bases de datos: la generación actual de DBMS proporcionan rutinas de comunicación especiales que permiten que la base de datos acepte las solicitudes del usuario final en un ambiente de red de computadoras. Por ejemplo, el DBMS podría disponer de funciones de comunicación para tener acceso a la base de datos a través de una interfaz web.

3.5 Metadatos

La primera acepción que se le dio al metadato (y actualmente la más extendida) fue la de *dato sobre el dato* (29). Aunque el término metadato no significa exactamente lo mismo en todos los campos (informáticos, bibliotecarios, dentro de una base de datos o en formato físico), en el entorno que compete a esta tesis, el entorno de los sistemas digitales, se usa cada vez más para referirse a los datos sobre los recursos de información. Un metadato puede incluir información descriptiva sobre el contexto, calidad, condición o características del recurso.

Los metadatos en sí no son algo nuevo. Lo que es nuevo es la multitud de formatos (tanto físicos como digitales) que se están diseñando y la forma de usarlos.

En el mundo bibliotecario el concepto de metadato, si bien no era empleado como se hace actualmente, existía. Los registros bibliográficos que se han creado a lo largo de muchos años son esencialmente metadatos.

Regresando a una definición más general, se debe tomar en cuenta que el metadato no sólo incluye descripción bibliográfica, sino que también contiene información relevante como materias implicadas, condiciones de uso, formato, etc.

Un metadato describe los atributos de un recurso, teniendo en cuenta que el recurso puede consistir en un objeto bibliográfico, inventarios, recursos audiovisuales, implementaciones de software, etc. Inclusive se podría decir que los metadatos pueden incluir todo lo que tenga relación con el dato que se describe (29), aunque esta definición podría generar problemas al momento de definir la estructura que se desea para los metadatos, es importante tomarla en cuenta para comprender que un metadato puede describir cualquiera de los atributos de un recurso u objeto.

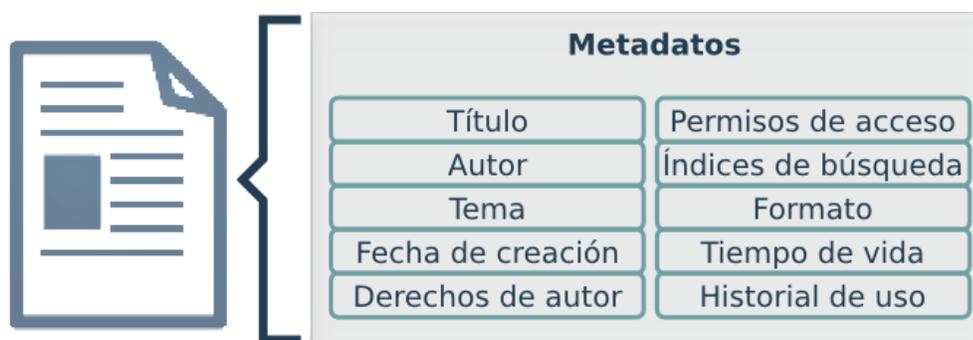


Figura 7 – Ejemplo de metadatos

Entonces, una definición formal y completa sería la siguiente:

«Metadato es toda aquella información descriptiva sobre el contexto, calidad, condición o características de un recurso que tiene la finalidad de facilitar su recuperación, autenticación, evaluación, preservación o interoperabilidad» (30).

Como se puede observar, un metadato describe a un recurso informático, pero también debe permitir ciertas funciones que va más allá de simplemente detallar un recurso.

Aunque la definición provista anteriormente agrupa en 5 campos las funciones que debe facilitar un conjunto de metadatos, es posible hacer un listado más riguroso acerca de las funciones que un metadato debe proveer o facilitar (31):

- Resumir el significado de los datos.
- Permitir la búsqueda.
- Determinar si el dato es el que se necesita.
- Proveer control de acceso.
- Recuperar y usar una copia del dato.
- Mostrar instrucciones de cómo interpretar un dato.
- Obtener información sobre las condiciones de uso (derechos de autor).
- Aportar información acerca de la vida del dato.
- Ofrecer información relativa al propietario o creador.
- Indicar relaciones con otros recursos.
- Controlar la gestión.
- Uso de ese recurso (historial del recurso).

Estos aspectos se pueden aplicar a la mayor parte de los metadatos y se pueden utilizar tanto en recursos digitalizados como en los no digitalizados. La idea que se halla al momento de definir funciones tan específicas es crear una estructura que permita al propio creador del documento asignar sus metadatos. En este sentido, los metadatos pueden crearse con la ayuda de herramientas para la generación automática de metadatos por un servicio de información, como podría ser una herramienta de software.

Por otro lado, debemos también enfocar especial atención en el *recurso* al cual describe un metadato, que puede ser tanto un documento como un simple dato, pero también puede ser una parte de un documento o un conjunto de datos. Dado que el presente trabajo se centra en información digital en distintos formatos (texto, imagen, video, audio, etc.), conviene integrar un nuevo elemento al hablar de metadatos: el objeto, como se entiende conceptualmente en el paradigma de la programación orientada a objetos (OOP, por sus siglas en inglés).

Un objeto es una entidad de información que es vista como un solo elemento y que, por tanto, es tratada de forma individual. Un objeto contiene atributos, los cuales son datos que conforman al objeto y lo representan de la manera que sea necesaria dentro del entorno en que se encuentran (32).

Cabe mencionar que esta definición es puramente conceptual y por tanto deja fuera conceptos técnicos propios de la implementación.

Entonces, desde un punto de vista más técnico, los recursos, como un documento (así como sus partes: líneas, párrafos e imágenes), pueden ser considerados como objetos, y los metadatos como los atributos que definen las características de cada uno de ellos, sin limitarse a su descripción.

Entonces, los metadatos dentro de un sistema tal como un acervo digital o un repositorio institucional, permitirían, no solo la descripción de los objetos sino también la indización de gran cantidad de datos de diferentes tipos sin necesidad de utilizar muchos recursos ya que se indiza la representación del objeto y no el objeto en sí. De este modo, un usuario que quisiera acceder a cierto recurso (objeto) analizaría tan solo sus metadatos para determinar cuál es el recurso específico que necesita, simplificando las tareas de búsqueda y recuperación, tanto para los usuarios como para el sistema subyacente. Tal esquema se ilustra en la **Figura 8**.

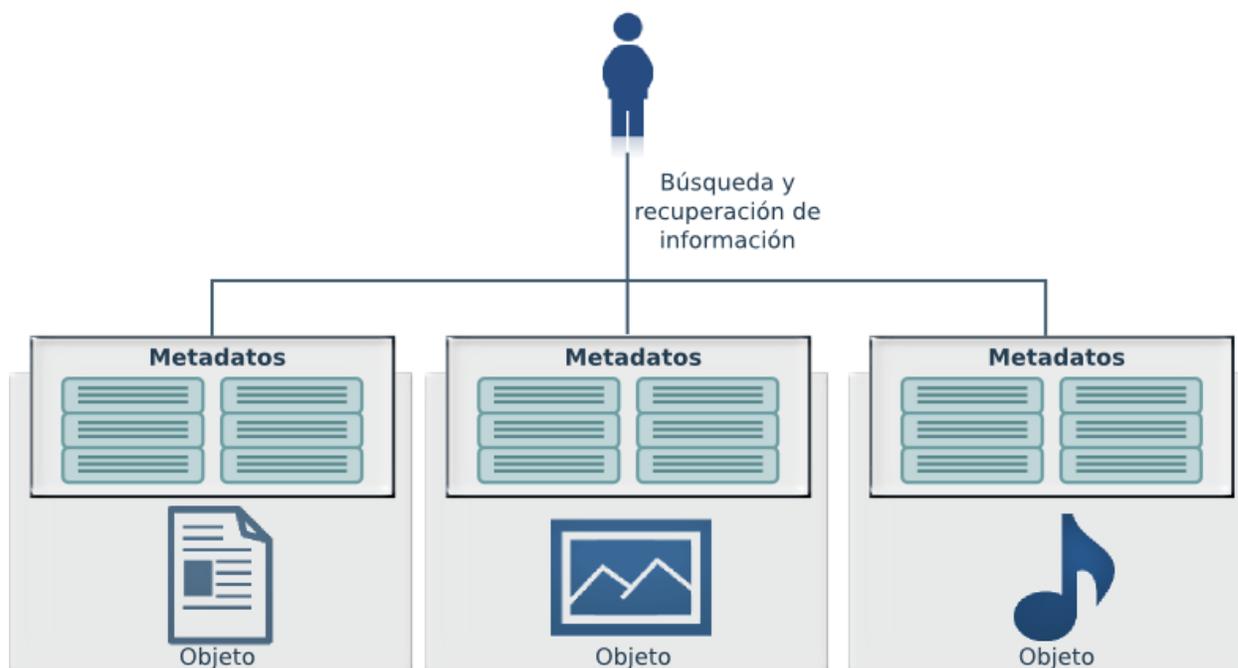


Figura 8 - Metadatos de objetos

3.5.1 Modelos de metadatos

Existe una gran variedad de modelos en la actualidad. Además, la mayoría de los sistemas digitales que utilizan metadatos para identificar sus objetos tienden a generar sus propios modelos.

Las instituciones, como las universidades y empresas, deciden generar sus propios modelos debido a los requerimientos sumamente específicos que sus proyectos contienen, aunado al poco interés que existe en generar un entorno de interoperabilidad, es decir, la mayoría de estos proyectos son sistemas aislados cuya información no pretende ser compartida con otros sistemas.

Sin embargo, este no es el caso de todos los acervos o repositorios digitales, existen una gran cantidad de proyectos que utilizan algún sistema de metadatos. Algunos de estos sistemas son:

- DC: Dublin Core Metadata.
- RDF: Resource Description Framework.
- TEI: Text Encoding Initiative.
- URC: Uniform Resource Characteristic/Citations.
- MARC DTD: MACHine Readable Cataloging Document Type Definition.
- GILS: Government Information Locator Service. Usado por los gobiernos de USA y Canadá.

Dublin Core es uno de los esquemas de metadatos más utilizados en la actualidad ^[5]. Se refiere al Dublin Metadata Core Element Set (Conjunto básico de elementos metadatos de Dublín), el cual es un proyecto cooperativo de ámbito internacional cuyo objetivo principal es crear un conjunto de metadatos que describan los documentos electrónicos de las redes con el fin de facilitar su búsqueda y recuperación.

Está formado por 15 elementos y su diseño busca ser intuitivo, para que los propios autores de la información puedan establecer los metadatos de sus documentos al mismo tiempo que los crean. Una característica también destacable es que estos elementos ponen más énfasis en apoyar a las características funcionales del sistema, como el acceso y recuperación de la información, que en la descripción del objeto.

Los elementos que lo componen son:

1. Título
2. Autor o creador
3. Materias y palabras clave
4. Descripción
5. Editor
6. Otras contribuciones
7. Fecha
8. Tipo de recurso
9. Formato
10. Identificador del recurso

⁵ Una lista actualizada de proyectos que voluntariamente se han registrado en Dublin Core Metadata Initiative se puede encontrar en <http://dublincore.org/projects/> (2 de agosto de 2012).

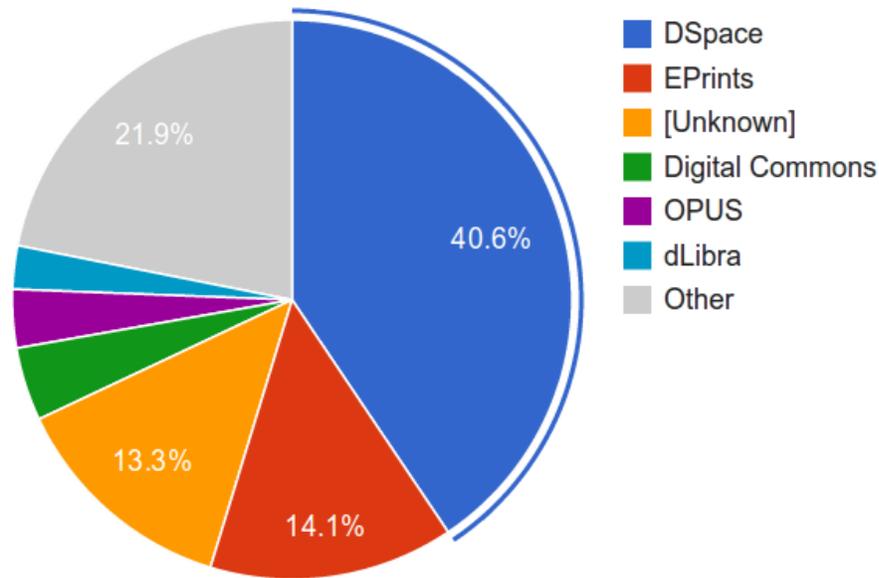
11. Fuente
12. Idioma
13. Relaciones
14. Alcance
15. Gestión de derechos

Un proyecto que utilice Dublin Core, generalmente estará mejor preparado para poder incluir interoperabilidad entre distintos repositorios o acervos, incluso si se están utilizando sistemas como MARC o alguno otro citado anteriormente, puesto que existen esquemas de conversión de metadatos para los modelos más usados, aunque la implementación de estos esquemas es responsabilidad de los desarrolladores.

Una gran cantidad de proyectos que utilizan DC lo hacen por medio de un CMS, que se encarga de las operaciones básicas para el manejo de contenido, incluyendo los metadatos. Uno de los CMS más ampliamente difundido, en la actualidad, es DSpace, el cual integra Dublin Core como modelo de metadatos ^[6].

La **Figura 9** muestra una comparativa entre los CMS más utilizados a nivel mundial dentro de un entorno de acceso abierto, es decir, son aquellos sistemas en donde los usuarios no requieren registrarse ni pagar cantidad alguna. DSpace conforma el 40.6% con un total de 890 proyectos.

⁶ El sitio oficial de Dspace <http://www.dspace.org/whos-using-dspace> (2 de agosto de 2012) lista todos los proyectos que se han registrado voluntariamente como usuarios de Dspace. La lista contiene más de 1300 proyectos.



Total = 2193 repositories

OpenDOAR - 17-Aug-2012

Figura 9 - CMS a nivel mundial (33)

La desventaja que existe al utilizar algún modelo existente de metadatos es que debe existir un proceso de adaptación. Se debe encontrar un balance entre el modelo que el proyecto requiere y el modelo que se pretende implementar. Es por ello que, como se mencionó al inicio de este capítulo (3.5.1), diversas instituciones prefieren generar sus propios modelos, puesto que alguno existente, como DC, podría carecer de la estructura específica que un sistema en particular necesita para satisfacer las necesidades de los usuarios.

3.6 Web 2.0

El término Web 2.0 es relativamente reciente. Apareció en el 2004 con una sesión de lluvia de ideas realizada entre las compañías O'Reilly (editorial enfocada en publicaciones de tecnología) y MediaLive International (empresa de medios de comunicación). La conferencia en donde se realizó la sesión estaba enfocada en la evolución de la Web. La **Figura 10** muestra algunos de los términos comúnmente asociados con Web 2.0.



Figura 10 – Conceptos de Web 2.0 [7]

En la siguiente conferencia sobre este tema (octubre 2005), Tim O'Reilly, presidente de la empresa O'Reilly, definió el concepto de Web 2.0, como «*el diseño de sistemas que logran que los efectos de la conectividad a redes hagan que la gente las use mejor, o lo que llamaríamos coloquialmente “potenciar la inteligencia colectiva”*» (34).

A partir de este momento, diversos autores han escrito definiciones sobre Web 2.0. Más de 7 años han pasado y todavía parece no existir alguna convención o acuerdo acerca de lo que significa Web 2.0, inclusive se habla de Web 1.5, Web 3.0, etc., para intentar darle cierta temporalidad a la Web.

Existen un par de definiciones que han sido muy difundidas y ampliamente aceptadas, no por ello significa que sean correctas, pero se expondrán a continuación para dar un panorama más amplio del vago concepto que es la Web 2.0:

«*Web 2.0 son todas aquellas utilidades y servicios de Internet que se sustentan en una base de datos, la cual puede ser modificada por los usuarios del servicio, ya sea en su contenido (añá-*

⁷ Imagen tomada de Wikipedia http://upload.wikimedia.org/wikipedia/commons/a/a7/Web_2.0_Map.svg (10 de agosto de 2012)

diendo, cambiando o borrando información o asociando datos a la información existente), bien en la forma de presentarlos, o en contenido y forma simultáneamente» (35).

«La Web 2.0 es una forma de entender Internet que, con la ayuda de nuevas herramientas y tecnologías de corte informático, promueve que la organización y el flujo de información dependan del comportamiento de las personas que acceden a ella [...]» (36).

Como se puede observar son definiciones que toman puntos de vista muy distintos, la primera hace hincapié en la parte técnica de la Web, colocando a las bases de datos como los pilares sobre los que se sostienen los servicios de una Web 2.0. La segunda definición, apoya más una actitud que una definición como tal. Pero ambas definiciones, junto con la original, dada por O'Reilly, concuerdan en un punto: la forma de interacción entre los usuarios y la información.

Entonces, podríamos entender que un servicio de Web 2.0 es aquel centrado en los usuarios, donde son ellos mismos quienes aumentan el valor del servicio por medio de su información. El usuario tiene cierto control sobre el sistema y puede agregar, modificar y eliminar datos del mismo, además de la capacidad de personalizar y configurar aspectos de su interés.

Existen aún muchas opiniones divergentes respecto a la Web 2.0, inclusive se opta por vetar el término, puesto que muchas veces ha sido usado injustificadamente para cualquier cosa a la que se le quiera dar un matiz de novedad. Pero enfocándonos en la acepción del párrafo anterior, podemos considerar sistemas tales como repositorios o acervos digitales, como aplicaciones Web 2.0, puesto que el conocimiento es propio de los usuarios y son ellos quienes aportarán documentos, imágenes, videos, etc. a los sistemas para beneficio de la institución o grupo al que pertenezcan.

**CONTEXTO Y ANÁLISIS DEL
PROBLEMA**

4.1 Situación actual del Departamento de Portugués en el CELE

Actualmente, existen pocos recursos didácticos de portugués en el CELE y éstos datan de los años 70's y 80's; están desactualizados y resultan poco interesantes o atractivos para los alumnos. Además, hay docentes que enseñan la variante europea del portugués y otros la variante brasileña.

El único material disponible en el Departamento de Portugués es el Curso Activo de Portugués (CAP), el cual está diseñado para trabajar bajo un enfoque comunicativo de enseñanza. Los recursos del CAP han caducado temáticamente, puesto que hacen referencia a eventos cuya trascendencia los convierte en materiales perecederos, al menos dentro de un contexto educativo (por ejemplo, noticias y artículos periódicos). Además, el CAP utiliza únicamente la variante brasileña del idioma portugués.

El Departamento de Portugués recientemente ha adoptado un nuevo enfoque de enseñanza, el cual está basado en *Tareas*. Cada tarea es una actividad en la que el alumno debe obtener un resultado específico que lo ayude a progresar en la comprensión y dicción del idioma portugués. Dependiendo del enfoque de cada tarea es posible clasificarlas en *Competencias*, los cuales describen una capacidad que el alumno puede desarrollar con respecto a la comprensión de una lengua.

El actual enfoque del Departamento incluye 47 tareas clasificadas en 6 Competencias:

1. Metalingüística
2. Metacognitiva
3. Estratégica
4. Discursiva
5. Sociolingüística
6. Lingüística

Bajo este nuevo enfoque, resulta complicado para el Departamento de Portugués seguir utilizado el CAP como medio para proveer de actividades a sus alumnos. Se requiere de una nueva colección de actividades que se apeguen al enfoque curricular basado en Tareas.

La propuesta del Departamento de Portugués incluye, no solo un nuevo enfoque, sino también la participación de sus docentes, quienes aportarán actividades que permitan el desarrollo de múltiples Competencias en el alumno. Es decir, serán los profesores quienes creen y utilicen todo el contenido que será disponible.

Para poder llevar a cabo una gestión óptima de los nuevos recursos que serán creados e integrar el control de calidad a los procesos de generación de contenido, la Didactiteca de Portugués requiere de un sistema que se encuentre completamente ligado a su enfoque curricular e implemente un control de calidad basado en la revisión por pares, descrita en el capítulo **3.1.4**.

La **Didactiteca de Portugués** es el sistema que dotará a los docentes de un lugar en el cual puedan aportar actividades y consultar los aportes de sus coetáneos. El flujo de trabajo dentro del proceso de revisión por pares será automatizado, de tal modo que no exista intermediario alguno, será el sistema quien se encargue de elegir a los dictaminadores y entregar las respuestas a todos los implicados en el proceso de revisión.

4.2 Uso e impacto de los acervos digitales en la educación

A lo largo de esta tesis se ha mencionado el impacto que tienen las TIC en la formación académica. Son muchos los beneficios que ha traído consigo el uso de la tecnología en el ámbito educacional, desde la inclusión de material audiovisual hasta la educación a distancia. Pero recientemente se han generado un conjunto de servicios que han logrado cambiar radicalmente el modo en que se enseña y se aprende. Estos servicios son los acervos digitales y repositorios institucionales. Se mencionan ambos términos debido a que suelen utilizarse para describir un mismo proyecto, aunque *repositorio* sigue siendo, por mucho, el término más popular (24).

Recordemos que la distinción de los dos términos mencionados ya ha sido descrita en el capítulo 3.3.2, pero esta distinción, lamentablemente, no puede ir más allá del ámbito teórico. En la práctica un repositorio y un acervo son homólogos, junto con catálogo, colección, archivo y, en algunos casos, biblioteca digital (22).

El interés por la correcta administración del capital intelectual ha existido, tanto en universidades como en la industria, desde hace mucho tiempo. Inclusive, las TIC se han utilizado para almacenar y gestionar contenido desde antes de la existencia de las bases de datos. Pero fue en el año 2002 cuando se empezó a tener un enorme interés por aplicar la tecnología en el ámbito educacional, haciendo uso de todos los recursos digitales con que se disponían. En este año, el Massachusetts Institute of Technology (MIT) ofreció libremente su catálogo de cursos en línea. El proyecto se conoce como *MIT OpenCourseWare (MIT-OCW)* e inició con 50 cursos en su catálogo, hasta llegar a 2100 cursos en el año 2012 (37).

El impacto del MIT provocó que otras universidades de prestigio internacional se unieran la iniciativa, generando proyectos propios como la *Red internacional de universidades Universia (OCW-Universia)* y *Multimedia Educational Resource for Learning and Online Teaching (MERLOT)*.

Para el análisis estadístico y comparativo se utilizaron dos fuentes ampliamente reconocidas, que proveen información catalogada y extensa acerca de los repositorios existentes. Estas fuentes son los proyectos de *Directory of Open Access Repositories (OpenDOAR)* (33) y *Registry of Open Access Repositories (ROAR)* (38).

Tanto OpenDOAR como ROAR listan solo aquellos proyectos que permitan el libre acceso a su contenido, es decir, un sistema que requiera de autenticación por parte de los usuarios o que solicite algún pago para poder obtener un servicio será excluido. El sistema que se generará como resultado de esta tesis no podría estar listado en ninguno de estos proyectos puesto que no permite ni el libre acceso ni la libre contribución. Los procesos de generación de contenido y acceso a los recursos deben pasar por uno o más filtros de seguridad.

El índice de OpenDOAR es el más extensa, con 2193 proyectos registrados a la fecha, pero solo refleja una porción de la verdadera cantidad de proyectos existentes para la preservación y

difusión de material en la red. Inclusive, algunos proyectos que cumplan con las características para ser aceptados en el índice de este sitio podrían no encontrarse debido a que el registro se realiza de forma manual; ni OpenDOAR ni ROAR son indizadores automáticos.

Tomando todo esto en cuenta, se analizarán algunos aspectos para conocer el uso y crecimiento de los proyectos de repositorios y acervos digitales.

A nivel mundial se contabiliza un total de 43 repositorios específicamente dedicados a recursos educativos abiertos (38), y además se puede encontrar este tipo de contenido en un 15.5% del total de los proyectos, lo cual corresponde a 342 proyectos (33). La **Figura 11** muestra los distintos tipos de contenido que los proyectos en OpenDOAR poseen. Como puede apreciarse, los artículos, tesis y tesinas se encuentran en la mayoría de los repositorios, acervos y bibliotecas digitales, mientras que los objetos de aprendizaje se encuentran por debajo, incluso, de las referencias bibliográficas.

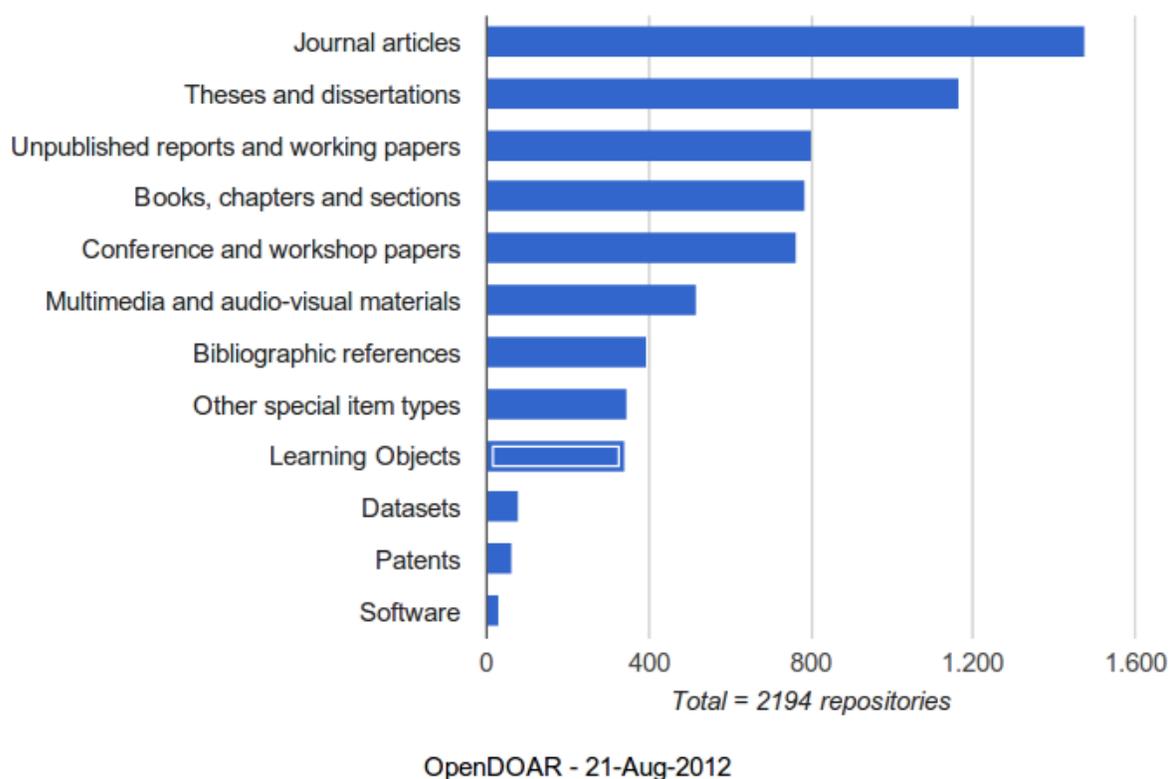


Figura 11 - Contenidos predominantes (33)

Las cifras anteriores ponen de manifiesto una característica significativa para los repositorios y acervos: son poco comunes los proyectos en donde se almacenen exclusivamente recursos educativos. La tendencia es a crear híbridos, con materiales de investigación y de docencia. Algunas instituciones han creado colecciones específicas para los contenidos de aprendizaje; otras han separado el repositorio educativo del resto de la producción científica. Finalmente hay institucio-

nes que participan en iniciativas OCW, dando acceso a cursos específicos pero sin abarcar el resto de recursos educativos depositados.

Existen múltiples explicaciones sobre este fenómeno, pero las dos más importantes se refieren a la reticencia de dos de los actores más importantes para cualquier proyecto enfocado en la enseñanza o aprendizaje: la institución formativa y los profesores (39).

Una institución no mantendrá un proyecto indefinidamente sin signos de éxito. El costo de la creación y gestión de los repositorios varía considerablemente pero, sea cual sea su costo total, es evidente que la inversión económica, material y humana que requieren los acervo o repositorios tendrá que estar justificada. El uso del contenido es la principal forma de valorar un proyecto de esta índole, tanto la cantidad de usuarios como una retroalimentación positiva son elementos clave para poder declarar un repositorio o acervo como una herramienta exitosa.

En cuanto a los profesores, los principales factores que contribuyen a la desconfianza de depositar los trabajos en los repositorios o acervos son muy variados (39):

- Desconocimiento de los conceptos teóricos y técnicos del sistema.
- Desconocimiento de los procedimientos para almacenar contenido.
- Falta de tiempo.
- Falta de medios.
- Indiferencia a los posibles beneficios.
- Resistencia a los cambios de procedimiento.
- Resistencia a un nuevo sistema.
- Miedo a la pérdida de control de su obra y al posible plagio.
- Desacuerdo con la idea de compartir su trabajo.

Entonces, nos vemos ante un panorama en donde el éxito de un acervo depende en gran parte de las personas involucradas y no solo en factores técnicos, puesto que la tecnología para la gestión de contenidos está lista para recibir proyectos de toda índole y extensión geográfica, tanto los sistemas ya existentes, como las plataformas para el desarrollo de sistemas propios.

Para combatir ambos problemas, muchas instituciones reconocen la necesidad de realizar esfuerzos de promoción (40). En el campo de la investigación, la difusión ha sido siempre un elemento básico. Es por ello que los repositorios cuyo contenido está conformado principalmente por artículos y resultados de investigaciones suelen ser los más divulgados, puesto que son las mismas investigaciones las que propician la promoción, tanto de los autores del contenido como del sistema que utilizan para publicar. En cambio, si los contenidos se conforman por material educativo, generalmente estarán dirigidos a coadyuvar en la enseñanza y el fortalecimiento de competencias docentes. La difusión dentro del contexto educativo, si bien es importante, no suele

tener la misma consideración que los factores relacionados directamente con la educación (calidad para la enseñanza, contexto cultural, temas de interés para los estudiantes, teorías del aprendizaje, etc.).

Para el caso de los proyectos con enfoque educativo, la promoción más efectiva suele ser la que llevan a cabo los autores que ya han comprobado las ventajas de utilizar cierto sistema, como por ejemplo el aumento de citas o las facilidades al buscar información especializada.

Un estudio a nivel mundial, realizado por investigadores de la Universidad de Southampton, informa que el 25.4% de los autores se informaron de algún repositorio o acervo a través de un colega, además, 8.5% recibió información de coautores de otras instituciones. Entre ambos se cuenta con el 33.9%, que representa a aquellas personas cuyo interés por los acervos y repositorios se vio influenciado por alguna persona. Este porcentaje supera el de aquellos que se informaron mediante los esfuerzos llevados a cabo por su institución, biblioteca o departamento que suma el 19.8%; donde 11.3% es por parte de una institución y 8.5% por parte de alguna escuela (41).

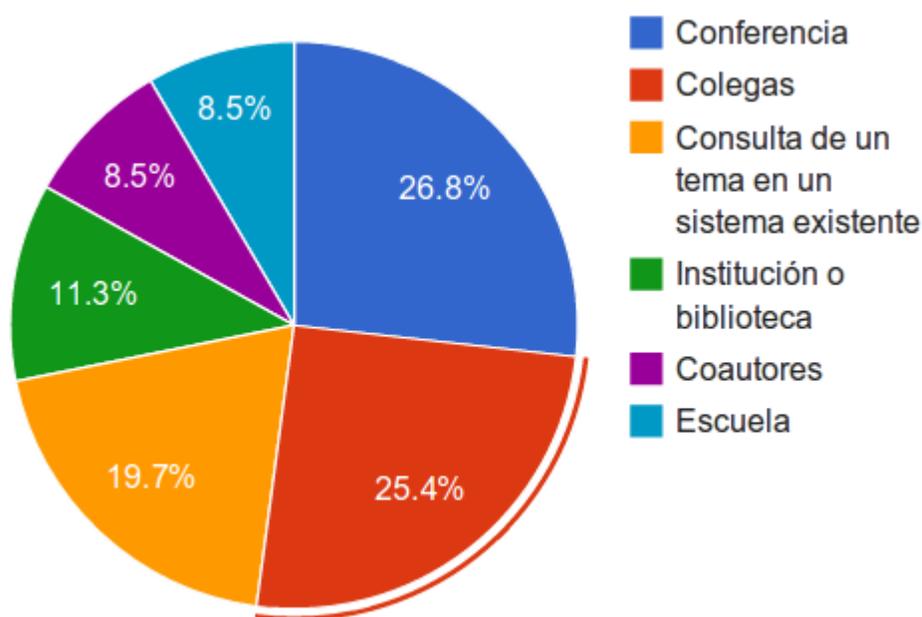


Figura 12 - Fuentes de información para decidir participar en un repositorio

La importancia de los acervos dentro de la comunidad educativa también se ha puesto de manifiesto con las numerosas iniciativas llevadas a cabo por la Unesco (42) y por otras entidades internacionales no gubernamentales, para asegurar el derecho universal a la educación y el aprendizaje a lo largo de la vida (39).

La **Figura 13** refleja el crecimiento de OpenDOAR desde finales del 2005 hasta los primeros meses del presente año. En el último mes del año 2005 existían tan solo 128 repositorios listados

en OpenDOAR. Lo cual nos demuestra el enorme auge que se ha tenido durante los últimos años (2193 proyectos registrados a la fecha).

Growth of the OpenDOAR Database - Worldwide

OpenDOAR - 17-Aug-2012



Figura 13 - Línea del tiempo OpenDOAR (33)

Para el caso particular de México, el listado se reduce enormemente, representa menos del 1% del total. OpenDOAR no provee una gráfica para todo el continente americano, por tanto la *Figura 14* muestra tan solo la sección de Norteamérica, donde México cuenta con 20 proyectos registrados, lo cual representa un 4.3% de esta sección del continente.

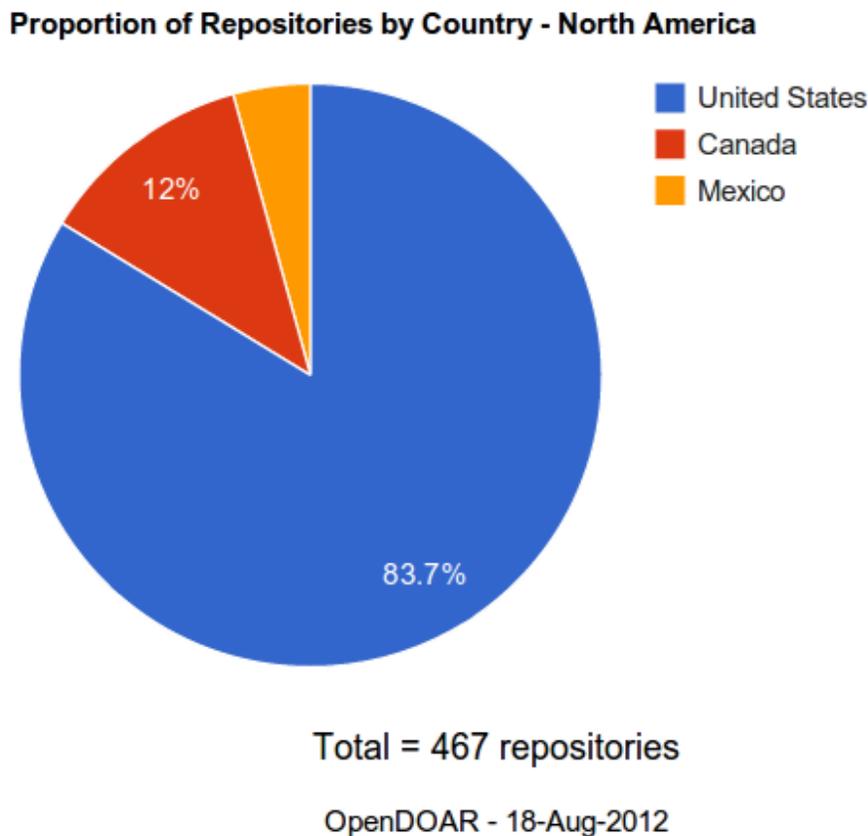


Figura 14 - Repositorios en Norteamérica (33)

Sin embargo, hay una tendencia interesante para nuestro país, que está relacionada con el tipo de contenido y no con la cantidad de proyectos existentes. La mayoría de los materiales actuales está estrechamente relacionada no solo con la difusión científica y cultural, sino también con la docencia. La **Figura 15** muestra una gráfica comparativa de todos los tipos de contenidos existentes, en donde las publicaciones periódicas, junto con las tesis y tesinas, son las predominantes. Pero a estos le siguen los contenidos bibliográficos, material multimedia y objetos de aprendizaje. Realizando la comparación con la **Figura 11** puede observarse que los objetos de aprendizaje tienen una mayor presencia en los proyectos mexicanos que en el promedio global.

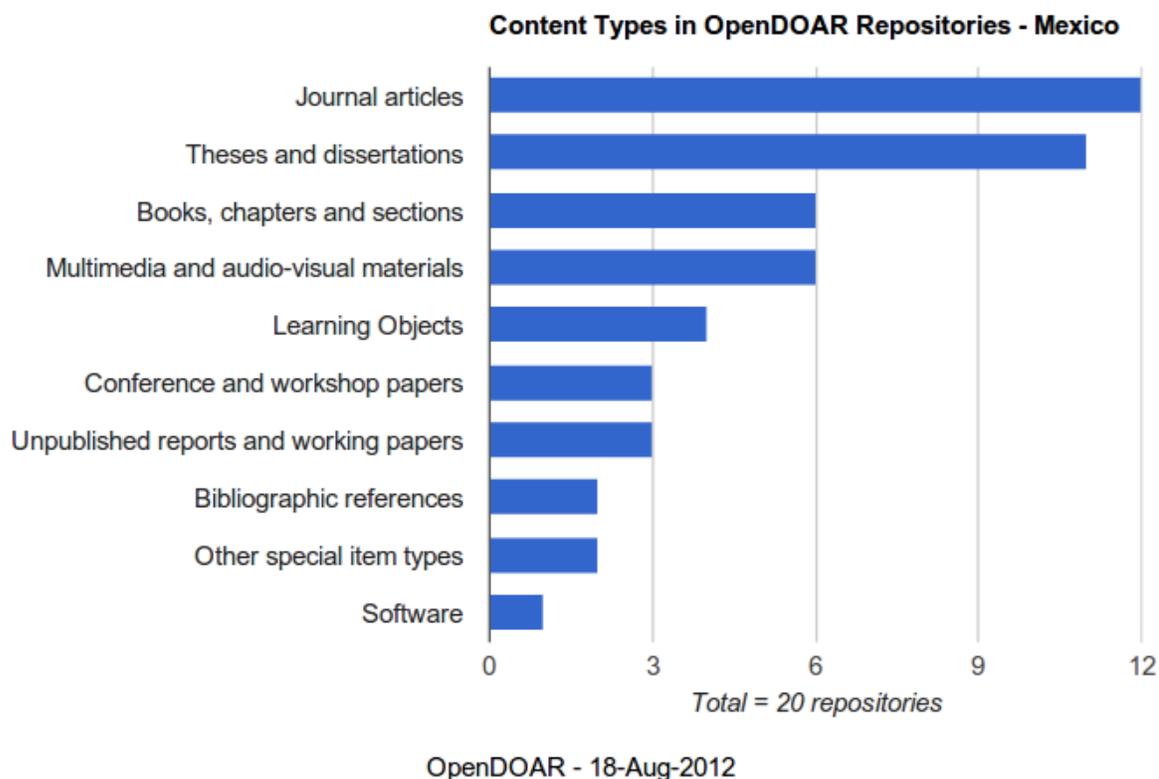


Figura 15 - Tipos de contenido en México (33)

Pero no solo es el tipo de contenido lo interesante para el caso de México, también lo es el tipo de institución u organización en donde se generaron los acervos. La **Tabla 3** enlista todos los proyectos mexicanos registrados en OpenDOAR.

Proyecto e institución	Recursos	Contenido
Acervo Digital del Instituto de Biología de la UNAM <i>Instituto de Biología, Universidad Nacional Autónoma de México</i>	12533	Multimedia
Artemisa en línea <i>Instituto Nacional de Salud Pública</i>	6818	Artículos
Biblioteca Sor Juana Inés de la Cruz <i>Universidad del Claustro de Sor Juana</i>	2776	Artículos Referencias Tesis
Centro de Recursos para la Enseñanza y el Aprendizaje <i>Universidad de Guadalajara</i>	202	Objetos de aprendizaje Multimedia Software
Colección de Tesis digitales <i>Centro Interactivo de Recursos de Información y Aprendizaje, Universidad de las Américas Puebla</i>	4223	Tesis
Coleccion de Tesis Digitales – UAEH <i>Universidad Autónoma del Estado de Hidalgo</i>	650	Tesis

Proyecto e institución	Recursos	Contenido
Colpos digital <i>Colegio de Postgraduados</i>	849	Tesis
Desarrolla, Aprende y Reutiliza <i>Tecnológico de Monterrey</i>	513	Objetos de aprendizaje Multimedia
Documentación en Ciencias de la Comunicación <i>Departamento de Estudios Socioculturales, ITESO</i> <i>Universidad Jesuita de Guadalajara</i>	5482	Artículos Tesis Libros
Dspace on Instituto Politécnico Nacional <i>Instituto Politécnico Nacional</i>	9841	Tesis
EduDoc, Centro de documentación sobre educación <i>ITESO Universidad Jesuita de Guadalajara</i>	488	Artículos Conferencias Tesis Libros Material sin publicar Objetos de aprendizaje
Publications of the Interactive and Cooperative Technologies Lab <i>Universidad de las Américas Puebla</i>	111	Artículos Referencias Material sin publicar
Redalyc <i>Universidad Autónoma del Estado de México</i>	243946	Artículos
Repositorio Académico Digital UANL <i>Universidad Autónoma de Nuevo León</i>	2389	Artículos Tesis
Repositorio de la Facultad de Filosofía y Letras <i>Facultad de Filosofía y Letras, Universidad Nacional</i> <i>Autónoma de México</i>	2253	Artículos Conferencias Libros Objetos de aprendizaje Multimedia
Repositorio Digital de la Universidad Veracruzana <i>Bibliotecas de la Universidad Veracruzana, Universidad</i> <i>Veracruzana</i>	11422	Artículos Tesis Libros Material especial
Repositorio Institucional- UNAM <i>Universidad Nacional Autónoma de México</i>	55392	Artículos Conferencias Libros Material sin publicar Multimedia
RU-Económicas <i>Instituto de Investigaciones Económicas, Universidad</i> <i>Nacional Autónoma de México</i>	1043	Artículos Tesis Libros Multimedia Material especial
Scientific Electronic Library Online México <i>Dirección General de Bibliotecas, Universidad Nacional</i> <i>Autónoma de México</i>	1069	Artículos

Tabla 3 - Repositorios en México

Como puede observarse, los proyectos pertenecen a organizaciones de carácter formativo, todas ellas de estudios superiores. El sector educativo pareciera ser el único interesado en la gestión de contenidos apoyada por las herramientas de TIC, obviando, por supuesto, el hecho de que la tabla contiene tan solo proyectos de libre acceso.

Existen proyectos cuyo propósito fundamental es la difusión, como lo pudiera ser un repositorio de artículos de investigación; tal es el caso de *Artemisa en línea* (43), cuyo contenido está compuesto en su totalidad por artículos de investigación de la salud.

Pero los proyectos que nos son de especial interés son aquellos cuyo objetivo va más allá de la difusión de sus materiales, son proyectos enfocados a la educación, que pretenden poner a disposición del público docente y estudiantil sus recursos. Un ejemplo un tanto ambiguo pudiera ser el caso de cualquier biblioteca digital, como la *Biblioteca Sor Juana Inés de la Cruz* (44) que provee a sus usuarios con material de carácter histórico, además de publicaciones actuales hechas por su profesorado. Cualquier persona, sea alumno, profesor o investigador puede interesarse en lo que esta biblioteca digital ofrece. Aun así, el propósito principal de la biblioteca no es la difusión de sus artículos de investigación, sino apoyar a los profesores y alumnos (en especial a aquellos que asisten a la universidad a la cual pertenece la biblioteca), como lo indica su página oficial.

Enfocándonos específicamente en aquellos proyectos totalmente dedicados a la educación, es posible filtrar de la lista de OpenDOAR tres proyectos mexicanos:

- Centro de Recursos para la Enseñanza y el Aprendizaje, CREA (45)
- Desarrolla, Aprende y Reutiliza, DAR (46)
- Centro de documentación sobre educación, EduDoc (47)

El sitio oficial de CREA, expone claramente el objetivo de su proyecto:

«El Centro de Recursos para la Enseñanza y el Aprendizaje (CREA) consiste en un repositorio de materiales educativos orientado al fortalecimiento del proceso de enseñanza-aprendizaje y a la formación integral de los estudiantes de nivel superior y medio superior. Los materiales educativos de CREA ofrecen la posibilidad de realizar simulación de situaciones reales para la solución de problemas, enriquecen la experiencia formativa de estudiantes y facilitan el proceso de enseñanza para los profesores, potencian la colaboración entre pares, y harán más eficiente el manejo de los recursos, pues su finalidad es ser reutilizados en diferentes contextos» (45).

Puede apreciarse que CREA está totalmente enfocado a la educación. Analizando la colección de materiales que posee, resulta evidente que va dirigido principalmente a los profesores, puesto que la temática predominante es llamada dentro de su clasificación como *educación*, la cual engloba estrategias de enseñanza, lineamientos de evaluación de cursos, manuales para el correcto uso de sistemas educativos en línea, técnicas grupales, etc. CREA tiene un objetivo que comparte ciertas similitudes con el trabajo de esta tesis, por su enfoque educativo y el énfasis en la instruc-

ción del personal. Sin embargo, contiene también documentos para los alumnos, como técnicas de estudio, administración del tiempo y resúmenes de conceptos teóricos de diversas áreas del conocimiento (matemáticas, ciencias biología, química, etc.)

DAR, por su parte se define como *«una iniciativa que aborda los desafíos inherentes de la formación de investigadores educativos. Busca apoyar la enseñanza e investigación mediante la creación de acceso a una colección creciente de acceso abierto de recursos y materiales digitales producidos y desarrollados por la comunidad académica»* (46).

A diferencia de CREA, cuyo contenido puede ser tanto para profesores como para alumnos, el proyecto DAR está dirigido exclusivamente a la docencia. Los temas que abarcan son:

- Administración del Conocimiento.
- Desarrollo de Competencias.
- Evaluación del Aprendizaje.
- Metodología para la investigación.
- Metodología y estrategias de enseñanza.
- Modelos innovadores de gestión educativa.
- Uso de las tecnologías en la educación.

Los cuales pretenden mantener en constante actualización al personal docente, tanto en las áreas puramente pedagógicas como en la innovación y uso de TIC.

Por último está EduDoc, cuyo objetivo es muy similar al de CREA, como puede apreciarse en la definición que el propio sitio del proyecto ofrece:

«EDU-DOC, es un centro de documentación [...] Este centro de documentación ofrece acceso sistematizado a reportes de investigación, material didáctico, ensayos, proyectos, ponencias y otros trabajos inéditos sobre el campo educativo, realizados por profesores, investigadores, estudiantes y profesionales de lo educativo.

Su propósito es propiciar la vinculación entre la producción académica sobre los temas educativos y los actores de la educación; brindar orientaciones y propuestas prácticas de aplicación en el quehacer cotidiano; y fomentar el debate de los temas de actualidad educativa para promover la circulación de las ideas y enriquecer la discusión teórica y práctica» (47).

EduDoc es también un proyecto enfocado en los profesores, enfatizando la importancia de mantener actualizados a sus usuarios al dotarlos de artículos e investigaciones recientes para que puedan ser puestas en práctica las más recientes técnicas pedagógicas. Al revisar el contenido de este acervo, es posible observar que el material contenido frecuentemente está relacionado con la niñez y adolescencia, lo cual indica que el acervo no está dedicado solo a educadores de nivel universitario, a pesar de ser un proyecto creado y mantenido por una universidad.

Además de los tres proyectos analizados, existe el Repositorio de la Facultad de Filosofía y Letras de la Universidad Nacional Autónoma de México (48), pero éste también se enfoca en la difusión de las investigaciones que la facultad realiza; además de contener una hemeroteca, donde la revista más antigua de la facultad, que puede consultarse libremente en formato digital, fue publicada en el año 1993. Es por ello que no es posible considerar a este repositorio como un proyecto completamente dedicado a la docencia, pero tampoco debe de excluirse, considerando que la cantidad de proyectos híbridos a nivel mundial es muy elevada.

El uso de los acervos y repositorios sigue en constante crecimiento, además, se observan actitudes positivas de los usuarios docentes respecto a la utilidad de los mismos, en donde indican que tales sistemas ayudan a preservar, reutilizar y crear nuevos materiales educativos. Otros usuarios indican que el uso de acervos y repositorios fomenta la motivación de los estudiantes y crear un vínculo con la institución. Por último algunos docentes consideran razones más altruistas al momento de compartir su conocimiento en un acervo, como el simple apoyo a la comunidad donde pertenece (40).

Para el caso específico de México, se vislumbra un panorama aún más alentador en cuanto al contexto educativo. En este país 20% de los proyectos incluyen material educativo, mientras que el promedio a nivel mundial es de 15.5% (33). De seguir con esta tendencia México incrementaría su presencia ante las comunidades docentes a nivel mundial, además de satisfacer la creciente necesidad de información, tanto de alumnos como profesores.

La Didactiteca de Portugués permitirá, precisamente, contar con un acervo que satisfaga las necesidades de sus usuarios docentes para proveer al alumnado de recursos actualizados y de calidad. Aun cuando el sistema no cumpla con las características idóneas para poder ser listando dentro de OpenDOAR o ROAR, su impacto no quedará limitado por este hecho, puesto que, como se ha analizado a lo largo de este capítulo, son los usuarios del sistema los que propician su crecimiento y difusión en un entorno educativo.

4.3 Análisis de acervos implementados en instituciones educativas

Para conocer cómo es que otras instituciones han abordado la realización de un proyecto con un enfoque primordialmente educativo se han escogido 5 acervos que son de especial interés:

- MIT OpenCourseWare (MIT OCW), Massachusetts Institute of Technology. (49)
- Multimedia Educational Resource for Learning and Online Teaching (MERLOT), California State University. (50)
- Centro de Recursos para la Enseñanza y el Aprendizaje (CREA), Universidad de Guadalajara. (45)
- Desarrolla, Aprende y Reutiliza (DAR), Tecnológico de Monterrey. (46)
- Funes: repositorio digital de documentos en Educación Matemática, Universidad de los Andes. (51)

Algunos de estos proyectos ya han sido mencionados anteriormente, como el MIT OCW por su enorme trascendencia o los proyectos CREA y DAR por estar desarrollados en México. Otros proyectos como Funes y MERLOT han sido escogidos específicamente para este análisis debido a que son ejemplos fundamentales cuyas características técnicas u objetivos son relevantes para el desarrollo de esta tesis.

El análisis considera los siguientes aspectos:

- Total de contenidos en el acervo
- Clasificación de los contenidos
- Formatos aceptados
- Software
 - Sistema de gestión de contenidos
 - Otras herramientas de gestión
 - Software para generación o edición de contenidos
- Esquema de metadatos
- Tipos de búsqueda
- Público objetivo
- Políticas de acceso a los contenidos
- Costo de acceso

- Políticas de envío de nuevo contenido
- Método de evaluación de nuevo contenido
- Licencias utilizadas

4.3.1 MIT OpenCourseWare

El proyecto MIT OCW ha sido tomado como un modelo a seguir por muchos proyectos a nivel mundial e inclusive se ha formado toda una red en donde las universidades generan su propio sistema OCW como apoyo a la iniciativa de libre acceso del MIT, para poner a disposición del mundo docente y estudiantil cursos de calidad impartidos por profesionales (49).

La información recabada pertenece al sitio oficial del MIT OCW (37) y a datos provistos por una de las administradoras del sistema (Shelly Upton, Web Production Specialist, comunicación personal, 23 de enero de 2012).

- Total de contenidos en el acervo: aproximadamente 2100 cursos, MIT OCW no ofrece una cantidad exacta.
- Clasificación de los contenidos
 - Aeronáutica y Astronáutica
 - Antropología
 - Arquitectura
 - Atletismo, Educación Física y Recreación
 - Ingeniería Biológica
 - Biología
 - Cerebro y Ciencias Cognitivas
 - Ingeniería Química
 - Química
 - Ingeniería Civil y Ambiental
 - Comparativa de estudios mediáticos
 - Ciencias de la Tierra, Atmósfera y Sistemas Planetarios
 - Ciencias económicas
 - Ingeniería Eléctrica y Ciencias de la Computación
 - División de Ingeniería de Sistemas
 - Grupo de Estudio Experimental
 - **Lenguas y Literaturas Extranjeras**
 - Ciencias de la Salud y Tecnología
 - Historia
 - Lingüística y Filosofía
 - Literatura
 - Ciencia de Materiales e Ingeniería
 - Matemáticas
 - Ingeniería Mecánica

- Artes y ciencias mediáticas
 - Artes Musicales y Teatrales
 - Ciencia e Ingeniería Nuclear
 - Física
 - Ciencias Políticas
 - Ciencia, Tecnología y Sociedad
 - Escuela de Administración Sloan
 - Programas Especiales
 - Recursos de consulta
 - Estudios y Planificación Urbana
 - Estudios de Género y de la Mujer
 - Estudios de Escritura y Humanísticos
- Formatos aceptados
 - Documentos: Portable document format (PDF)
 - Imágenes: estándar Joint Photographic Experts Group (JPG, JPEG)
 - Videos: Windows Media Video (WMV), RealMedia Video (RM, RV, RMVB), Audio Video Interleave (AVI) y Flash Video (FLV) para streaming en el sitio.
 - Software
 - Sistema de gestión de contenidos: Plone, sistema de código abierto escrito en el lenguaje de programación Python.
 - Otras herramientas de gestión: EdgeSuite (sistema de distribución de contenidos), FileMaker Pro (aplicación de BD relacionales)
 - Software para generación o edición de contenidos: el sistema no es partícipe en el proceso de creación de contenido.
 - Esquema de metadatos: esquema diseñado específicamente para el MIT OCW, posee compatibilidad con SCROM, el cual es un estándar enfocado en objetos pedagógicos.
 - Tipos de búsqueda:
 - Búsqueda simple: Palabras clave
 - Búsqueda avanzada:
 - Palabras clave con modificadores para buscar frase exacta y excluir palabras.
 - Secciones en que se divide un curso.
 - Opción para elegir la cantidad de resultados a desplegar

- Opción para el ordenamiento, ya sea por relevancia o por fecha.
- Público objetivo: profesores, estudiantes y autodidactas a nivel mundial.
- Políticas de acceso a los contenidos: El contenido de MIT OCW está disponible para todo el público, sin previo registro o aprobación de los administradores.
- Costo de acceso: gratuito
- Políticas de envío de nuevo contenido: Solo los docentes del MIT pueden agregar nuevo contenido, el proceso es interno y se desconoce su funcionamiento. Esta política se asemeja al modo en que trabajará la Didactiteca de Portugués.
- Método de evaluación de nuevo contenido: Revisión por pares.
- Licencias utilizadas: todos los contenidos poseen la misma licencia para su libre acceso y distribución, la cual es Creative Commons Attribution-NonCommercial-ShareAlike 3.0 United States (CC BY-NC-SA 3.0)^[8].

4.3.2 Multimedia Educational Resource for Learning and Online Teaching

El proyecto de MERLOT no almacena ningún contenido, simplemente sirve como un catálogo que reúne material de diversos sitios. MERLOT almacena únicamente direcciones URL. Pero esto no es limitante para incluir al proyecto dentro del análisis, puesto que es uno de los repositorios más antiguos que siguen en funcionamiento, al haber sido creado en 1997.

La información recabada pertenece al sitio oficial del MERLOT (50) y a datos provistos por uno de los administradores del sistema (MERLOT Webmaster, comunicación personal, 23 de enero de 2012).

- Total de contenidos en el acervo: 38137
- Clasificación de los contenidos:
 - Colecciones
 - Arte
 - Negocios
 - **Educación**
 - Humanidades
 - Matemáticas y Estadística
 - Ciencia y Tecnología

⁸ La versión actual de esta licencia puede consultarse en el sitio oficial de Creative Commons: <http://creativecommons.org/licenses/by-nc-sa/3.0/us/> (Consultado: 28 de agosto de 2012)

- Ciencias Sociales
- Desarrollo Laboral
- Tipos de material
 - Animación
 - Herramienta de evaluación
 - Tarea
 - Caso de estudio
 - Colección
 - Herramientas de desarrollo
 - Ejercicios y prácticas
 - ePortfolio
 - Repositorio de Objetos de Aprendizaje
 - Curso en Línea
 - Artículo
 - Libro de texto
 - Presentación
 - Pruebas y cuestionarios
 - Material de referencia
 - Simulación
 - Herramientas de Red Social
 - Tutorial
 - Taller y Material de formación
- Formatos aceptados: Potencialmente es cualquier formato, puesto que el recurso está hospedado en un sitio externo a MERLOT.
- Software
 - Sistema de gestión de contenidos: No utiliza algún sistema de gestión existente. El proyecto se desarrolló dentro de la California State University.
 - Otras herramientas de gestión: TopLink, plataforma de Java para almacenar objetos en una base de datos relacional.
 - Software para generación o edición de contenidos: el sistema no es partícipe en el proceso de creación de contenido.
- Esquema de metadatos: Esquema propio de MERLOT, no se ofrece especificación alguna.
- Tipos de búsqueda:
 - Búsqueda simple: Título, palabras clave y descripción.
 - Búsqueda avanzada de material:

- Palabras clave con opción para buscar por frase exacta
- Título
- URL del contenido
- Descripción
- Categorías y subcategorías
- Idioma
- Tipo de material: Animación, tarea, presentación, simulación, curso, artículo, etc.
- Formato del material: PDF, Flash, HTML, Word, Excel, etc.
- Compatibilidad con algún dispositivo móvil específico.
- Uso de alguna plataforma educativa: Moodle, eCollege, WebCT, etc.
- Autor
- Licencia
- Costo
- Fecha en que fue agregado a MERLOT
- Presencia de comentarios y revisiones por pares.
- Permite el ordenamiento por valoración, relevancia, título, autor, tipo de material y fecha.
- Búsqueda avanzada de miembros:
 - Palabras clave
 - Nombre
 - Apellido
 - Correo electrónico
 - Afiliación
 - Dirección
 - Tipo de miembro: estudiante, profesor, administrador, etc.
 - Fecha de registro
 - Cantidad de contribuciones

- Permite el ordenamiento por cantidad de contribuciones, relevancia, apellido, fecha, afiliación y tipo de miembro.
- Búsqueda en otros sitios: incluye MIT OCW, Wikipedia, Youtube, Google Books, IEEE Digital Library, etc.
- Público objetivo: Principalmente a profesores, pero el material es tan diverso que cualquiera interesado en los temas que MERLOT contiene puede encontrarlos de utilidad.
- Políticas de acceso a los contenidos: El contenido de MERLOT está disponible para todo el público, sin previo registro o aprobación de los administradores.
- Costo de acceso: el acceso a MERLOT es gratuito, pero algunos contenidos tienen el costo que el autor imponga en su sitio web.
- Políticas de envío de nuevo contenido: Todo miembro de MERLOT podrá contribuir con nuevo contenido. Para ello, el usuario deberá registrarse en el sistema.
- Método de evaluación de nuevo contenido: Revisión por pares, aunque la revisión se realiza después de haber sido publicado el contenido, puesto que se utiliza únicamente como una medida adicional de evaluación. Es decir, no todo contenido en MERLOT es revisado.
- Licencias utilizadas:
 - Derechos reservados, especificados por el autor del contenido.
 - Libre distribución, licencias Creative Commons en varias versiones.

4.3.3 Centro de Recursos para la Enseñanza y el Aprendizaje

CREA será analizado por ser un proyecto mexicano enfocado en la educación. Los acervos y repositorios mexicanos no son numerosos, al menos no aquellos listados en índices mundiales como OpenDOAR y ROAR, por ello es de gran importancia conocer lo que México está ofreciendo actualmente a sus profesores por medio de contenidos digitales.

La información recabada pertenece al sitio oficial del CREA (45)

- Total de contenidos en el acervo: 202
- Clasificación de los contenidos:
 - Formación
 - Estrategias de Enseñanza
 - Evaluación del Aprendizaje

- Gestión de la información
- Metodología para la investigación
- Recursos para el Aprendizaje
- Uso de Tecnologías de la Información y la Comunicación
- Áreas del conocimiento
 - Ciencias naturales y matemáticas
 - Ciencias sociales
 - Filosofía y psicología
 - Generalidades
 - Geografía e historia
 - Artes, bellas artes y artes decorativas
 - Lenguas
 - Literatura y retórica
 - Religión
 - Tecnología (Ciencias aplicadas)
- Formatos aceptados
 - Multimedia: Shockwave Flash (SWF), presentación de Microsoft PowerPoint (PPS, PPT, PPTX).
 - Ejecutables: Archivo ejecutable de Windows (EXE).
 - Otros: Formato de almacenamiento sin pérdida ZIP (ZIP).
- Software
 - Sistema de gestión de contenidos: DSpace.
 - Otras herramientas de gestión: ninguna.
 - Software para generación o edición de contenidos: Software de autoría de objetos de aprendizaje creador, desarrollado específicamente para el proyecto CREA. Esta herramienta restringe el formato final que tendrán los recursos.
- Esquema de metadatos: Dublin Core (utilizado por DSpace)
- Tipos de búsqueda:
 - Búsqueda simple: Título y palabras clave

- Búsqueda avanzada: se permite un máximo de 3 categorías
 - Palabras clave
 - Autor
 - Título
 - Tema
 - Resumen
 - Colección
 - Patrocinador
 - Identificador
 - Lengua
- Público objetivo: estudiantes y profesores a nivel mundial.
- Políticas de acceso a los contenidos: El contenido de CREA está disponible para todo el público, sin previo registro o aprobación de los administradores.
- Costo de acceso: gratuito
- Políticas de envío de nuevo contenido: Todo miembro de CREA podrá contribuir con nuevo contenido. Para ello, el usuario deberá registrarse en el sistema.
- Método de evaluación de nuevo contenido: revisión por parte del comité técnico.
- Licencias utilizadas: todos los contenidos poseen la misma licencia para su libre acceso y distribución, la cual es Creative Commons Atribución-NoComercial-CompartirIgual 2.5 México (CC BY-NC-SA 2.5)⁹

4.3.4 Desarrolla, Aprende y Reutiliza

DAR es otro proyecto mexicano enfocado en la educación, donde los materiales son elaborados por la comunidad docente. Es decir, DAR es un proyecto ideado por profesores y para profesores, lo cual lo convierte en un ejemplo de sumo interés para el presente trabajo.

La información recabada pertenece al sitio oficial del DAR (46).

- Total de contenidos en el acervo: 513
- Clasificación de los contenidos:
 - Administración del Conocimiento

⁹ La versión actual de esta licencia puede consultarse en el sitio oficial de Creative Commons: <http://creativecommons.org/licenses/by-nc-sa/2.5/mx/> (Consultado: 28 de agosto de 2012).

- Desarrollo de Competencias
- Evaluación del Aprendizaje
- Metodología para la investigación
- Metodología y estrategias de enseñanza
- Modelos innovadores de gestión educativa
- Uso de las tecnologías en la educación
- Formatos aceptados: el sitio de CREA parece aceptar potencialmente cualquier tipo de formato sin embargo solo los siguientes han podido ser encontrados en el acervo.
 - Documentos: Portable document format (PDF), documento de Microsoft Word (DOC, DOCX), XML de Microsoft Word (XML)
 - Imágenes: estándar Joint Photographic Experts Group (JPG, JPEG)
 - Multimedia: presentación de Microsoft PowerPoint (PPS, PPT, PPTX).
- Software
 - Sistema de gestión de contenidos: DSpace
 - Otras herramientas de gestión: ninguna
 - Software para generación o edición de contenidos: el sistema no es partícipe en el proceso de creación de contenido.
- Esquema de metadatos: Dublin Core (utilizado por DSpace)
- Tipos de búsqueda:
 - Búsqueda simple: Título y palabras clave
 - Búsqueda avanzada: se permite un máximo de 3 categorías
 - Palabras clave
 - Autor
 - Título
 - Tema
 - Resumen
 - Series
 - Patrocinador
 - Identificador

- Lengua
 - Público objetivo: profesores e investigadores docentes de todo el mundo.
 - Políticas de acceso a los contenidos: El contenido de DAR está disponible para todo el público, sin previo registro o aprobación de los administradores.
 - Costo de acceso: gratuito
 - Políticas de envío de nuevo contenido: Todo miembro de DAR podrá contribuir con nuevo contenido. Para ello, el usuario deberá registrarse en el sistema.
 - Método de evaluación de nuevo contenido: desconocida
 - Licencias utilizadas:
 - Derechos reservados, especificados por el autor del contenido.
 - Libre distribución, licencias Creative Commons en varias versiones.

4.3.5 Funes: repositorio digital de documentos en Educación Matemática

Funes es un proyecto muy peculiar, no ha tenido difusión alguna y no se encuentra listado en OpenDOAR o ROAR, a pesar de permitir el libre acceso a los contenidos. Aun así, el sitio registra una cantidad considerable de visitas y descargas, el último reporte disponible muestra más de 10000 visitas y 12000 descargas durante el mes de marzo del 2011 (52). En resumen, un proyecto creado en el año 2009 y puesto al público sin ninguna publicidad o difusión por parte de la institución creadora, tiene ahora más de 300 visitas diarias.

La explicación de este fenómeno es debido a que Funes ofrece a su comunidad información que no era posible obtener de otro modo, permite a los educadores matemáticos tener acceso a información actual, pero sobre todo, en su idioma. Funes es el único repositorio que ofrece libremente material de investigación en educación matemática en español.

La información recabada pertenece al sitio oficial del Funes (51) y a datos provistos por uno de los administradores del sistema (Pedro Gómez, Departamento de Didáctica de la Matemática, comunicación personal, 23 de enero de 2012).

- Total de contenidos en el acervo: 1259
- Clasificación de los contenidos:
 - Enfoque
 - Actividad
 - Ensayo
 - Innovación

- Investigación
- Otro enfoque
- Nivel Educativo
 - Educación Infantil (0-6 años)
 - Educación Primaria (7-12 años)
 - Educación Secundaria y Bachillerato (13-18 años)
 - Educación de adultos
 - Estudios de posgrado
 - Formación Profesional
 - Todos los niveles educativos
 - Título de grado universitario
 - Ningún nivel educativo
 - Otro nivel Educativo
- Formatos aceptados:
 - Documentos: Portable document format (PDF)
- Software
 - Sistema de gestión de contenidos: Eprints^[10], sistema de código abierto escrito en el lenguaje de programación Perl^[11].
 - Otras herramientas de gestión: ninguna.
 - Software para generación o edición de contenidos: el sistema no es partícipe en el proceso de creación de contenido.
- Esquema de metadatos: DublinCore (utilizado por Eprints)
- Tipos de búsqueda:
 - Búsqueda rápida: Título, palabras clave y resumen.
 - Búsqueda simple:
 - Título. palabras clave, resumen.
 - Permite ordenar por año, autor y título.

¹⁰ <http://www.eprints.org/>

¹¹ <http://www.perl.org/>

- Búsqueda avanzada:
 - Palabras clave
 - Título
 - Autor
 - Resumen
 - Fecha
 - Tipo de registro
 - Proyectos
 - Editorial
 - Editores
 - Estado
 - ID de registro
 - Permite ordenar por año, autor y título
- Público objetivo: La comunidad iberoamericana de educadores matemáticos.
- Políticas de acceso a los contenidos: El contenido de Funes está disponible para todo el público, sin previo registro o aprobación de los administradores.
- Costo de acceso: gratuito.
- Políticas de envío de nuevo contenido: Todo miembro de Funes podrá contribuir con nuevo contenido. Para ello, el usuario deberá registrarse en el sistema.
- Método de evaluación de nuevo contenido: revisión por pares.
- Licencias utilizadas: licencias Creative Commons en varias versiones.

PROPUESTA

5.1 Resultados de la comparativa entre los acervos seleccionados

Algunos de los datos analizados en el capítulo 4.3 son útiles para conocer el estado en que se encuentra un proyecto, tal como lo es conocer la cantidad de contenidos presentes, otros datos nos permiten analizar cómo y a quiénes están ofreciendo sus servicios. Sin embargo, no todos los aspectos analizados pueden ser considerados como variables al momento de diseñar un acervo, puesto que muchos de ellos son inherentes al tipo de herramienta que se desee crear.

Aun así, es posible llegar a conclusiones significativas al respecto. Por lo tanto, estos datos de carácter contextual fueron también analizados detenidamente:

- Con respecto a la cantidad de contenidos, los proyectos iberoamericanos parecen pequeños en comparación con MIT OCW y MERLOT, aunque no por ello menos importantes o de menor impacto.
- La mayoría de los proyectos pretenden alcanzar un público a escala global, solo Funes define un público objetivo más reducido y especializado.
- Crear clasificaciones extensas pareciera no ser un problema, puesto que es una característica en común entre los acervos analizados, aunque resulta mucho más rápido y sencillo utilizar las herramientas de búsqueda que navegar entre listados arborescentes de categorías y subcategorías.
- Todos los acervos analizados proveen libre acceso a los contenidos, aunque MERLOT no puede garantizar un uso completamente gratuito, puesto que no tiene ningún control sobre los sitios externos en donde se alojan los contenidos como tal.
- En ningún acervo se permite la libre publicación de nuevo contenido, siempre es necesario registrarse para poder realizar aportes. Con excepción del MIT OCW que no permite siquiera registrarse, puesto que solo los profesores del MIT tienen permitido publicar.
- La revisión por pares (revisión por expertos), es el sistema más ampliamente utilizado para garantizar la calidad del contenido que se pretende publicar. Aunque se desconoce cómo se realice la evaluación para cada uno de los acervos.
- Todos los sistemas incluyen manejo de licencias, en algunos casos es obligatorio usar una licencia en específico, mientras que otros acervos permiten a los autores elegir o definir sus propios términos. Sea cual sea el caso, la relevancia de los derechos de autor está siempre presente.

Por otro lado, existen aspectos técnicos, que permiten analizar qué clase de servicios son los ofrecidos por un acervo digital, ya sea uno ampliamente difundido, como el MIT OCW, o uno menos divulgado como Funes.

Para simplificar los resultados de la comparativa técnica, se presentarán algunas tablas con información digerida y separada por cada aspecto técnico de relevancia.

Primero, se analizó uno de los elementos de mayor impacto: el sistema de gestión de contenidos. La elección de éste puede significar un enorme cambio en todos los demás aspectos técnicos, puesto que los CMS existentes cuentan con sus propias limitaciones y estándares a los que un desarrollador debe apegarse. La tabla **Tabla 4** muestra los CMS junto con las demás herramientas de gestión que cada proyecto utiliza.

	MIT OCW	MERLOT	CREA	DAR	Funes
CMS	Plone + Desarrollo propio	Desarrollo propio	DSpace	Dspace	Eprints
Herramientas de gestión	EgdeSuite, FileMaker Pro	TopLink	Ninguno	Ninguno	Ninguno
Software para la generación o edición de contenido	Ninguno	Ninguno	Desarrollo propio	Ninguno	Ninguno

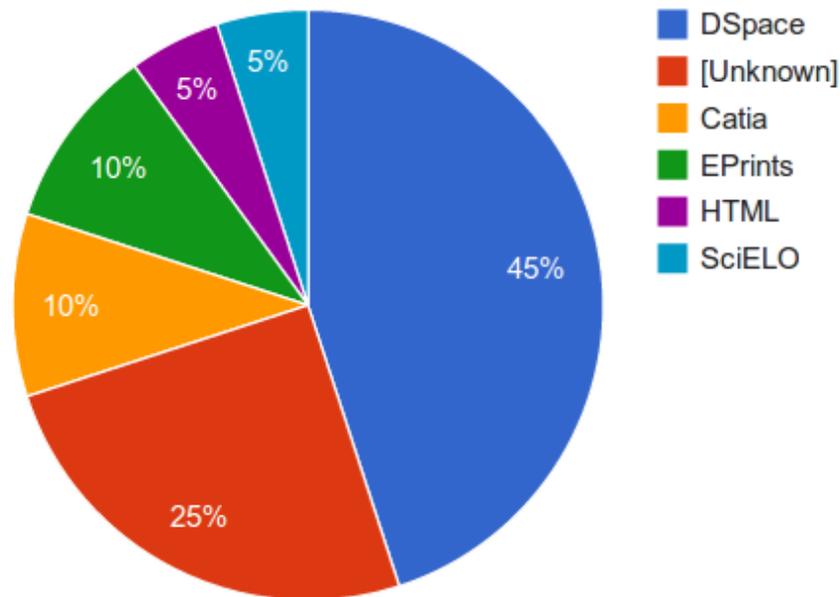
Tabla 4 - CMS en los acervos analizados

MERLOT es el único proyecto que no hace uso de un CMS existente, mientras que los proyectos mexicanos utilizan DSpace, el CMS más ampliamente difundido en México (33) (38).

El MIT OCW utiliza Plone, pero se sabe que, además, han contratado a un grupo de desarrolladores para realizar una gran cantidad de modificaciones al sistema (Shelly Upton, Web Production Specialist, comunicación personal, 23 de enero de 2012). Actualmente, el sistema sigue siendo actualizado y modificado por este grupo de desarrolladores, según se vaya necesitando. Es por ello que, si bien se está utilizando un sistema de gestión existente, el grupo de desarrollo del MIT ha generado gran parte del proyecto.

Además, existe un aspecto interesante para el caso específico de México. La **Figura 16** muestra una gráfica donde se agrupan los proyectos mexicanos por el tipo de CMS que esté utilizando. El 45% pertenece a Dspace, mientras que el 25% que se muestra como *desconocido*, pertenece a los desarrollos propios.

Usage of Open Access Repository Software - Mexico



Total = 20 repositories

OpenDOAR - 25-Aug-2012

Figura 16 - Uso de CMS en México (33)

Pero este porcentaje no representa fielmente a todas aquellas plataformas desarrolladas dentro de la misma institución. *Catia* (53), CMS listado en la gráfica, es un desarrollo propio cuyo uso se extiende únicamente a proyectos creados por la misma universidad a la que *Catia* pertenece, y así lo es también el proyecto que cita a HTML como un CMS, el cual es un sistema creado por los desarrolladores del acervo, perteneciente a *Publications of the Interactive and Cooperative Technologies Lab* (54).

Con lo cual, una gráfica que agrupe a todos los CMS desarrollados por la institución a la cual pertenece el acervo se puede observar en la **Figura 17**.

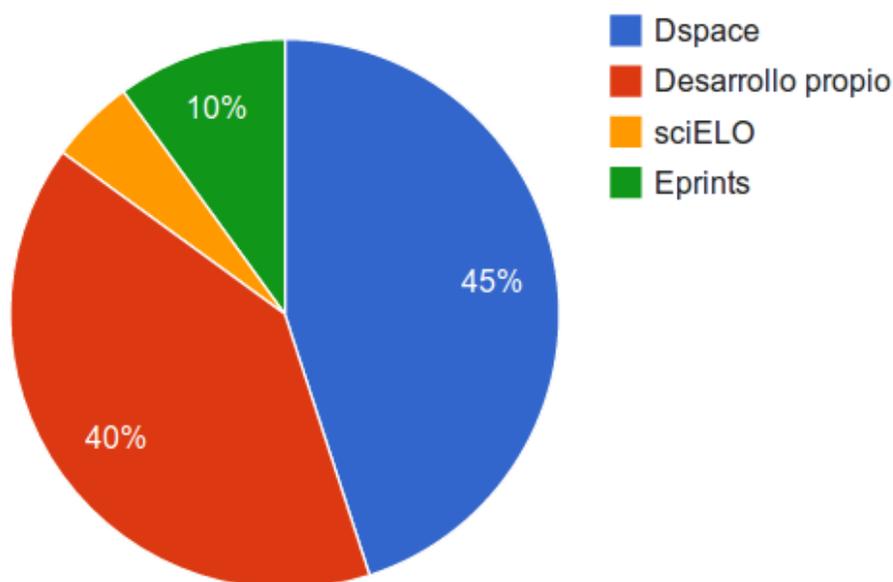


Figura 17 - CMS en México agrupando los desarrollos propios

Una gran parte de los acervos mexicanos utilizan un sistema propio, 40% de ellos, aun cuando Dspace sigue siendo la plataforma más ampliamente utilizada a nivel mundial.

Estos resultados permiten considerar ampliamente el realizar un sistema propio para la Didáctica de Portugués, en caso de que los CMS existentes no cumplan con las características necesarias o dispongan de restricciones que mellen el funcionamiento del sistema.

El segundo aspecto analizado está relacionado con el tipo de materiales que se permiten en un acervo, junto con los formatos aceptados. La *Tabla 5* muestra todos los formatos que pudieron ser reconocidos en los proyectos analizados.

	MIT OCW	MERLOT	CREA	DAR	Funes
Documentos	PDF	Cualquiera	Ninguno	PDF, DOC, DOCX, XML	PDF
Imágenes	JPG, JPEG	Cualquiera	Ninguno	JPG, JPEG	Ninguno
Video	WMV, RM, RV, RMVB, AVI, FLV	Cualquiera	Ninguno	Ninguno	Ninguno
Multimedia	Ninguno	Cualquiera	SWF, PPS, PPT, PPTX	PPS, PPT, PPTX	Ninguno
Otros	Ninguno	Cualquiera	EXE, ZIP	Ninguno	Ninguno

Tabla 5 - Tipos de materiales y sus formatos

La diversidad de formatos es enorme si consideramos que estos son tan solo 5 acervos de los miles existentes en diversas instituciones. De los proyectos analizados, CREA es el único que contiene archivos ejecutables y en formato de compresión, debido a que son archivos creados por su herramienta de apoyo para la generación de contenido (45). Además no existe un tipo de archivo que sea manejado por todos los acervos, lo cual indica que tanto el tipo de archivo como el formato están ligados a otras restricciones ajenas al acervo, como las herramientas que los autores poseen o el tipo de contenido que se necesita. Por ejemplo, el MIT OCW requiere de clases grabadas en video y por tanto soporta una gran variedad de formatos para ello, tanto para descarga como para la reproducción en el mismo sitio web, mientras que Funes requiere solo de documentos PDF, puesto que la información que contiene se compone únicamente de texto e imágenes embebidas en los mismos documentos.

El siguiente aspecto analizado fue el esquema de metadatos que cada acervo emplea. La **Tabla 6** resume esta comparativa, en donde puede apreciarse que solo existen dos variantes:

- Un esquema propio
- DublinCore

Además, cabe resaltar que aquellos sistemas que utilizan DublinCore lo hacen en parte porque su CMS, DSpace, tiene configurado ese esquema de metadatos como el principal a usar. DublinCore también ofrece significativas ventajas de interoperabilidad al conectar distintos acervos.

MIT OCW	MERLOT	CREA	DAR	Funes
Esquema propio	Esquema propio	DublinCore	DublinCore	DublinCore

Tabla 6 - Esquemas de metadatos

Entonces, el resultado de este análisis indica que para un desarrollo propio es común utilizar también un esquema de metadatos propio, pero si se busca la interoperabilidad, la mejor opción es utilizar un esquema estandarizado o en su defecto implementar un esquema propio y crear un proceso de interoperabilidad que traduzca el esquema a algún estándar (tal es el caso del MIT OCW) (37).

Por último, se analizaron los tipos de búsqueda que cada acervo incluye. La cantidad de términos que cada búsqueda contiene es totalmente dependiente de los metadatos que contenga el acervo y de cómo esté configurado el sistema. Por tanto, la *Tabla 7* muestra únicamente que tipos de búsqueda que ofrece cada sitio, además de mencionar si permite o no el ordenamiento de los resultados; cuando no se permite el ordenamiento, se entiende que los resultados son mostrados por orden de relevancia con los parámetros de la búsqueda realizada.

	MIT OCW	MERLOT	CREA	DAR	Funes
Búsqueda simple	✓	✓	✓	✓	✓
Búsqueda avanzada	✓	✓	✓	✓	✓
Permite ordenamiento de resultados	✓	✓	×	×	✓

Tabla 7 - Tipos de búsqueda

La búsqueda es una de las herramientas básicas de todo acervo o repositorio. Cuando se tiene una enorme cantidad de información, en ocasiones, tenerla estructurada por categorías no es suficiente, el usuario debe tener algún modo de especificar qué es lo que desea obtener. Es por ello que todos los acervos analizados poseen tanto búsqueda simple, la cual elige un grupo de metadatos en los cuales buscar, como búsqueda avanzada, en donde el usuario puede especificar con mayor detalle lo que desea buscar aplicando algunos filtros.

Opcionalmente 3 de los acervos permiten al usuario elegir el orden de los resultados, aunque por lo general una persona deseará que el primer elemento de la lista sea aquel que mejor se ajusta a sus necesidades y, por tanto, a su búsqueda; aun así es una opción que debe ser considerada.

5.2 Enfoque

Para la realización del proyecto se ha seguido el estándar de ingeniería de software *ISO/IEC 29110 Lifecycle profiles for Very Small Entities (VSEs)* (55). Este estándar está diseñado para grupos de trabajo reducidos, llamados *Entidades Muy Pequeñas* (VSEs, por sus siglas en inglés). Cada VSE conforma una empresa, departamento o proyecto de hasta 25 personas, aunque esto no significa que un grupo mayor no tenga motivos para seguir este estándar.

ISO/IEC 29110 proporciona los procesos de *Gestión de Proyecto e Implementación de Software*, los cuales integran prácticas basadas en la selección de elementos estándares de dos documentos:

- ISO/IEC 12207 Software life cycle processes
- ISO/IEC 15289 Content of systems and software life cycle process documentation.

Se ha elegido la ISO/IEC 29110 puesto que está destinada para establecer procesos que implementen cualquier enfoque o metodología de desarrollo, pero sin dejar a un lado las necesidades de la VSE o del proyecto.

Mediante el uso de la ISO/IEC 29110 se pueden obtener diversos beneficios (56), algunos de los más importantes son:

- La obtención de un conjunto acordado de requisitos para un proyecto, junto con los productos que el cliente espera recibir.
- Un proceso de gestión disciplinado que proporciona una completa visualización del proyecto, así como acciones correctivas para los problemas y desviaciones.
- Un proceso sistemático de implementación del software que satisfaga las necesidades del cliente y asegure la calidad del producto.

El propio documento de la norma explica detalladamente la logística, procesos y tareas que deben ser realizados para generar un producto de calidad y acorde con los requerimientos. Por lo tanto, a continuación se explicará únicamente un panorama general de los procesos más significativos que se han seguido en la realización de la Didactiteca de Portugués.

Los procesos básicos que existen al momento de desarrollar un proyecto de software, con el estándar ISO/IEC 29110, son dos:

1. Gestión del Proyecto
2. Implementación del Software

El propósito del proceso Gestión del Proyecto es establecer y llevar a cabo de manera sistemática las tareas del proyecto de implementación de Software, para poder cumplir con los objetivos del proyecto en calidad, tiempo y costo esperados.

El propósito del proceso Implementación de Software es la realización sistemática de las actividades de análisis, diseño, construcción, integración y pruebas para productos de software nuevos o modificados de acuerdo con los requisitos especificados.

La interrelación entre ambos procesos puede observarse en la **Figura 18**.

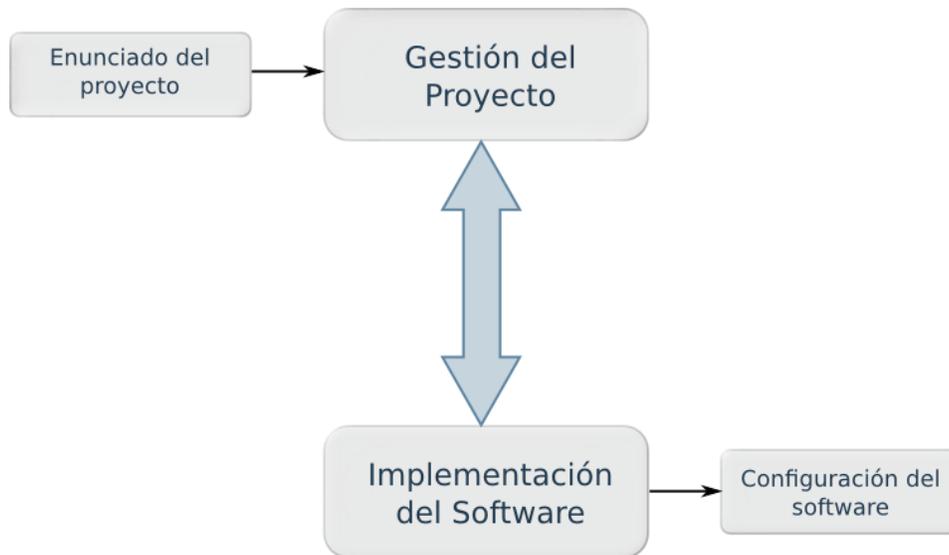


Figura 18 - Procesos básicos

5.2.1 Proceso de Gestión del Proyecto

El Proceso de Gestión del Proyecto consiste en las siguientes actividades:

1. Planificación del Proyecto
2. Ejecución del Plan del Proyecto
3. Evaluación y Control del Proyecto
4. Cierre del Proyecto

Planificación del Proyecto

Esta etapa se genera el Plan de Proyecto basándose en el Enunciado de Trabajo, el cual es un documento en donde cliente expone sus necesidades, junto con el producto que espera recibir. La planificación del proyecto incluye las siguientes tareas:

1. Revisión del Enunciado de Trabajo para analizar las tareas subsecuentes a realizar.
2. Establecer el ciclo de vida del proyecto, incluyendo las tareas a realizar y su duración.
3. Crear una estrategia para asegurar la calidad del proyecto, a través de la verificación y validación de los productos que le serán entregados al cliente.
4. Asignación de roles y responsabilidades

5. Se establecen las necesidades de entrenamiento del equipo, así como el costo y cronograma del proyecto.
6. Se identifican los riesgos del proyecto.
7. Creación de una estrategia para el control de versiones, el cual es la gestión de los cambios sobre los elementos de los productos.
8. Creación del repositorio del proyecto, para almacenar, gestionar y entregar de manera controlada los productos y versiones de documentos.

Ejecución del Plan del Proyecto

Una vez que el plan del proyecto ha sido definido, se implementa en la ejecución del proyecto. Entre las actividades que se deben realizar en esta etapa están:

1. Monitoreo de la ejecución del plan del proyecto
2. Análisis y evaluación del impacto, tanto en costo como tiempo, de cualquier solicitud de cambio, ya sea interno (por parte del equipo de trabajo) o externo (por parte del cliente).
3. Realizar reuniones con el equipo de trabajo para detectar problemas y revisar el estado del proyecto.
4. Realizar reuniones con el cliente para registrar acuerdos y dar seguimiento hasta la conclusión del proyecto.
5. Realizar respaldos periódicos del repositorio de acuerdo con la estrategia de control de versiones.
6. En caso de ser necesario, realizar recuperaciones del repositorio.

Evaluación y Control del Proyecto

La actividad de evaluación y control, permiten determinar el desempeño del plan contra los compromisos ya documentados. Dentro de esta etapa existen las siguientes actividades:

1. Evaluación del progreso del plan contra los objetivos planteados.
2. Realización de pruebas para comparar contra los resultados esperados.
3. Identificación de las desviaciones junto con su coste e impacto.
4. Identificar nuevos riesgos potenciales.
5. Adoptar acciones correctivas para los problemas detectados.

Cierre del Proyecto

En la etapa final de la gestión del proyecto se proporciona la documentación y los productos al cliente, de acuerdo con los requerimientos del contrato. Dentro de esta etapa se realizan las actividades que darán culminación al proyecto:

1. Entregar los productos al cliente, tal como se especificaron en la documentación.
2. Generar el Acta de Aceptación con el cliente, para formalizar la conclusión del proyecto, obteniendo todas las firmas necesarias.
3. Incluir en el repositorio del proyecto los resultados y el Acta de Aceptación.

El diagrama de la **Figura 19** muestra el proceso completo de Gestión del Proyecto.

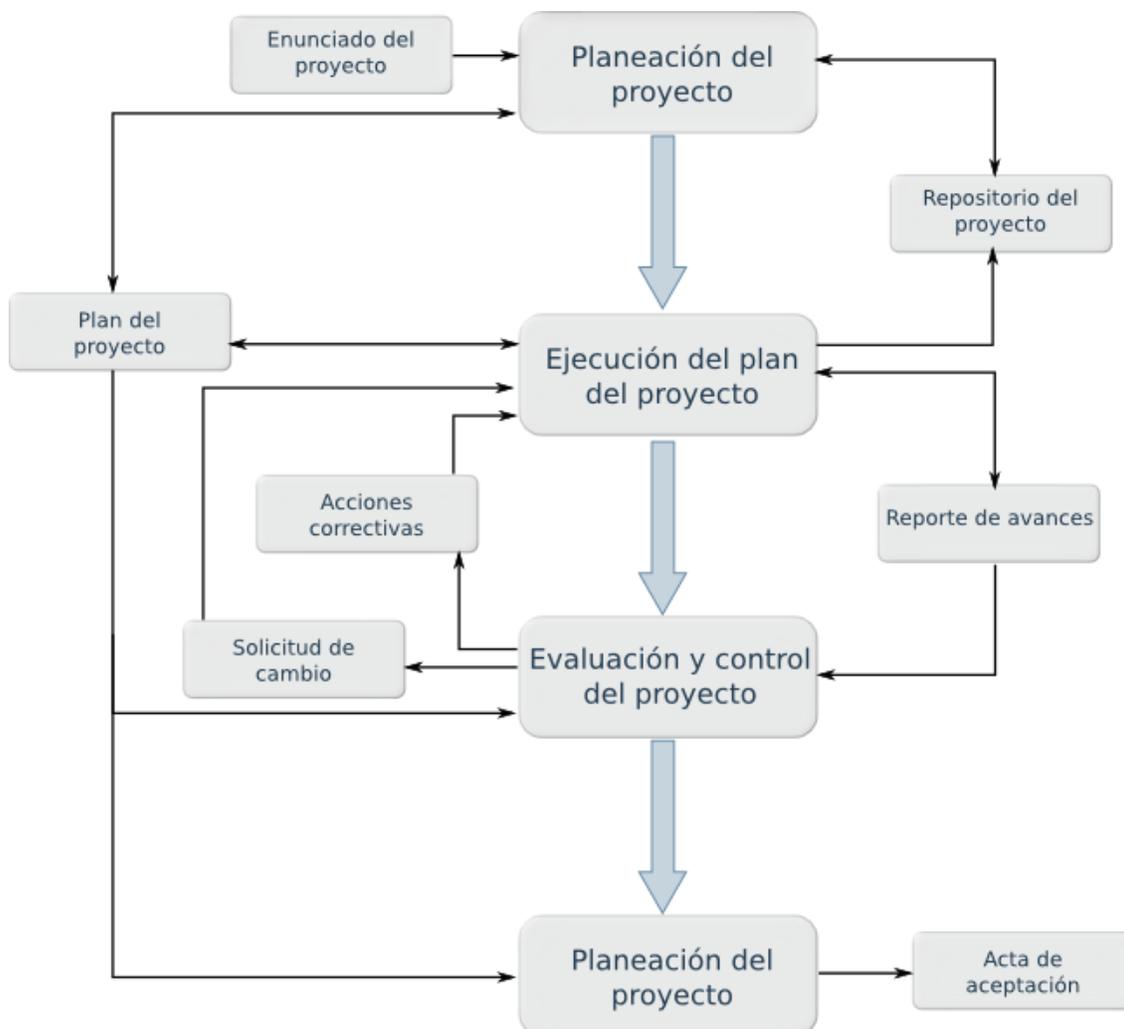


Figura 19 - Proceso de Gestión del Proyecto

5.2.2 Proceso de Implementación del Software

El Proceso de Implementación de Software consiste en las siguientes actividades:

1. Inicio de la Implementación del Software.
2. Análisis de Requerimientos del Software.
3. Arquitectura y Diseño Detallado del Software.
4. Construcción del Software.
5. Integración y Pruebas del Software.
6. Entrega del Producto.

Inicio de la Implementación del Software

La actividad Inicio de la Implementación del Software asegura que el Plan del Proyecto establecido en la actividad Planeación del Proyecto se lleve a cabo. La actividad incluye:

1. La revisión del Plan del Proyecto por parte del Equipo de Trabajo para determinar la asignación de las Tareas.
2. El establecimiento de un ambiente para la implementación.

Análisis de Requerimientos del Software

Durante esta fase de proceso se analizan y establecen los requisitos acordados con el Cliente, mediante la realización de las siguientes tareas:

1. Documentar o actualizar la Especificación de Requerimientos, consultando, además, otras fuentes de información, como proyectos anteriores, para obtener nuevos requerimientos.
2. Revisar del Plan del Proyecto por parte del Equipo de Trabajo para determinar la asignación de las Tareas.
3. Verificación y validar los requisitos.
4. Implementar el control de versiones de los requisitos del producto de software.

Arquitectura y Diseño Detallado del Software

La actividad de Arquitectura y Diseño Detallado del Software utiliza los requisitos de software para la creación de la arquitectura y diseño del sistema. Esta actividad incluye:

1. La asignación de tareas
2. El diseño de la arquitectura, los componentes de software y las interfaces asociadas.
3. El diseño detallado de los componentes de software y sus interfaces.

4. La revisión de los requerimientos por parte del equipo de trabajo.
5. La creación de los casos de prueba y procedimientos de verificación para las pruebas de integración.

Construcción del Software

En esta etapa del proceso se desarrollan el código y los datos de software, para ello se realizan las siguientes actividades:

1. Analizar y comprender el diseño de software.
2. Construir los componentes de software basándose en el diseño de software.
3. Diseñar o actualizar los casos de pruebas.

Integración y Pruebas del Software

La etapa de Integración y Pruebas del Software asegura que los componentes de software integrados satisfacen los requisitos del Software, para ello se realizan las siguientes actividades:

1. Establecimiento del ambiente de prueba.
2. Integración del software con los casos de prueba.
3. Realización de pruebas utilizando los casos de prueba diseñados en la etapa previa.
4. Se corrigen los defectos encontrados.

Entrega del Producto

Al final del proceso se provee el producto de software al cliente. La actividad incluye:

1. La configuración del software.
2. La realización de los documentos y manuales necesarios para el cliente.
3. Realizar la entrega tal como se acordó con el cliente.

La **Figura 20** muestra el proceso completo de Implementación del Software.

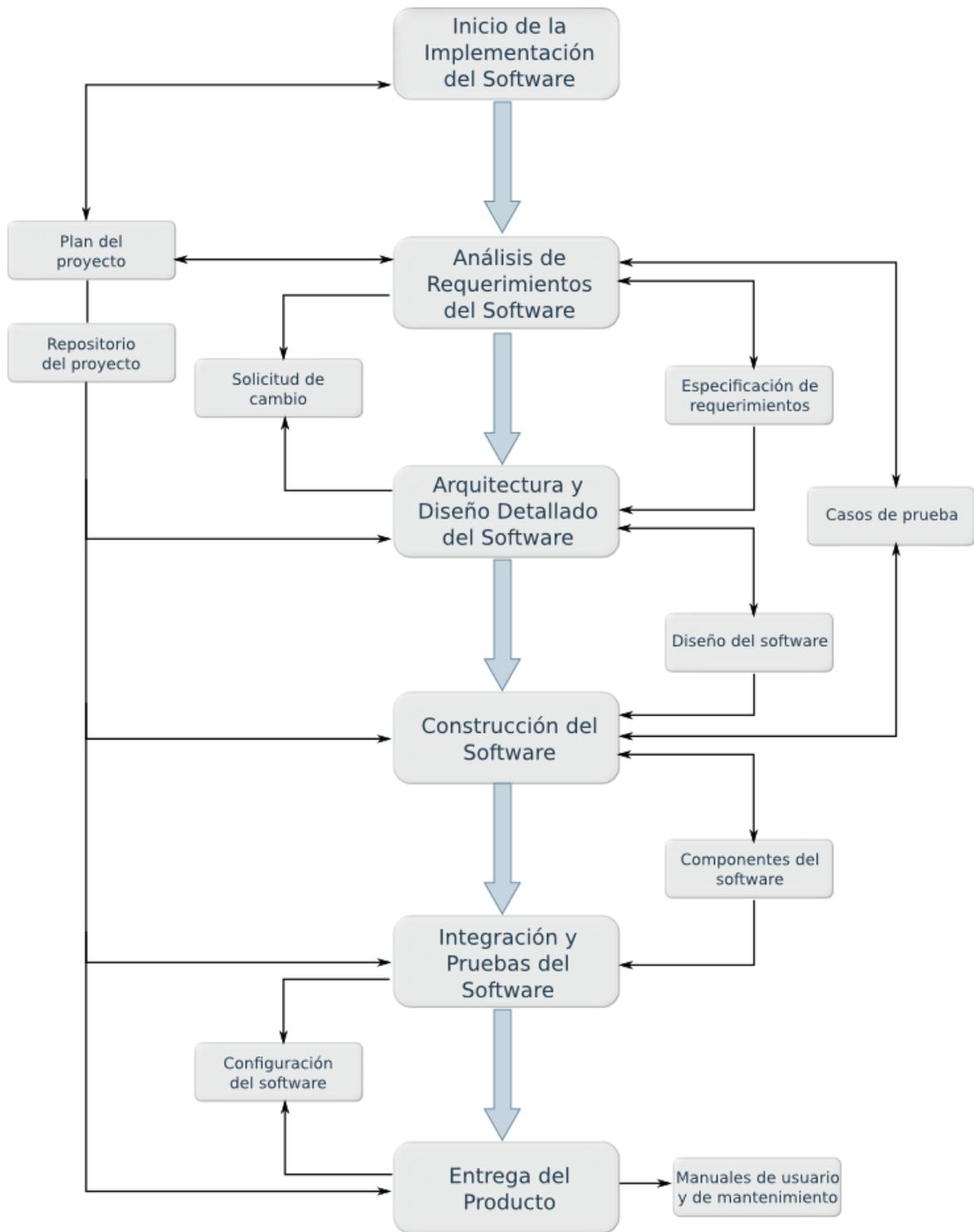


Figura 20 - Proceso de Implementación del Software

5.3 Diseño

Para diseñar el sistema se ha utilizado el estándar Unified Modeling Language (UML)^[12], el cual es un lenguaje que permite documentar de manera gráfica el análisis y diseño de un sistema de software. UML está diseñando, primordialmente, para un sistema orientado a objetos y su metodología más distintiva es conocida como *modelado de casos de uso* (57).

Un modelo de caso de uso describe las acciones de un sistema sin describir cómo realiza tales acciones. Además éste divide la funcionalidad de un sistema en acciones significativas para los usuarios del sistema, tales acciones son los casos de uso.

Siguiendo el proceso descrito en el capítulo 5.2, en específico la parte que compete a la implementación, se obtuvieron todos los requerimientos del proyecto. Los cuales se han dividido y estratificado para su comprensión y diseño. Del mismo modo se han categorizado a los usuarios del sistema dependiendo del rol que desempeñan.

Los requerimientos de la Didactiteca de Portugués han sido organizados en casos de uso y los usuarios en actores, los cuales representan los distintos roles que existen en el sistema.

Los actores existentes son:

- Público: Es cualquier persona que no se encuentra registrada en la Didactiteca de portugués.
- Miembro: Es una persona que se encuentra registrada en la Didactiteca y puede utilizar sus recursos. El Miembro es quien aporta nuevo material al acervo.
- Dictaminador: Es la persona encargada de evaluar el material que se agrega y elimina de la Didactiteca. Evalúa el material nuevo que un Miembro desea ingresar al acervo y el material desactualizado que un Administrador desea remover del acervo.
- Administrador: Es la persona encargada de revisar y gestionar los movimientos que se realizan en la Didactiteca. Publica el material aprobado por un Dictaminador y revisa periódicamente el acervo para eliminar material que no esté vigente o que haya caído en desuso. Tiene cierto control sobre el estado de las cuentas de los Miembros y decide los eventos de interés académico que serán mostrados en la Didactiteca.

Los casos de uso identificados se enlistan a continuación, donde algunos pueden englobar varias acciones las cuales han sido estratificadas para simplificar la representación del sistema:

1. Listar contenido
2. Buscar contenido
 - 2.1. Buscar de forma simple

¹² <http://www.uml.org/>

- 2.2. Buscar de forma avanzada
3. Visualizar parcialmente el contenido
4. Registrarse como miembro
5. Autenticar
6. Visualizar completamente el contenido
 - 6.1. Descargar contenido
 - 6.2. Generar PDF
7. Solicitar envío de contenido
8. Listar material de su autoría
9. Dictaminar
 - 9.1. Evaluar contenido y correcciones
 - 9.2. Evaluar eliminación de contenido
10. Administrar usuarios
 - 10.1. Modificar estado de cuenta
 - 10.2. Eliminar miembro
11. Administrar contenido
 - 11.1. Publicar material
 - 11.2. Eliminar material
12. Administrar eventos
13. Enviar contenidos a RU-CELE
14. Obtener listas de correo
15. Ver reportes
 - 15.1. Ver descargas
 - 15.2. Ver accesos

La **Figura 21** muestra el diagrama completo de casos de uso, en donde cada actor interactúa con un conjunto definido de casos de uso, dependiendo de las actividades que se le permita realizar dentro del sistema, aunado a los permisos que posea. Por ejemplo, el Público es el actor con menos permisos y, por tanto, menos acciones, ya que no está registrado en el sitio solo se le permiten vistas previas del contenido. En cambio el Administrador tiene mayor control sobre el sistema puesto que es el actor con más privilegios, aunque también con mayores responsabilidades.

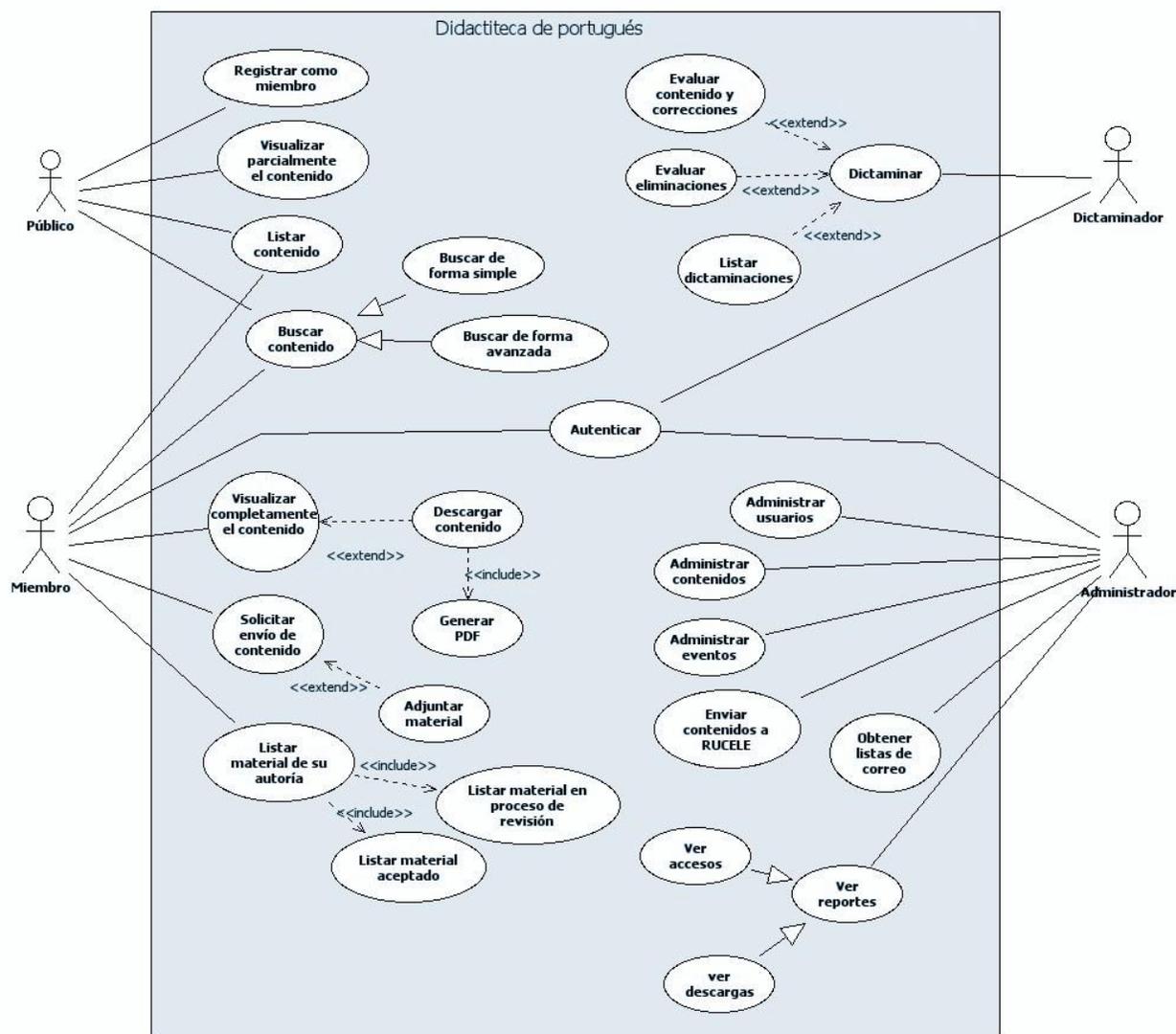


Figura 21 - Diagrama de casos de uso de la Didactiteca

Para poder implementar los casos de uso se dividió la aplicación en paquetes, cada uno de los cuales cumple con una tarea específica. La segregación en paquetes se hizo tomando en cuenta el tipo de procesos implicados y los recursos sobre los cuales se necesita realizar alguna acción. De este modo, todas aquellas acciones que se realicen sobre un recurso específico estarán englobadas en un solo paquete; por ejemplo, todo lo concerniente al contenido de la Didactiteca será realizado en el paquete Manejo de Contenido.

Los paquetes tienen también interacciones entre ellos dependiendo del proceso que sea necesario seguir para un caso de uso en particular; por ejemplo, todo proceso que requiera el ingreso de datos por parte de un usuario requerirá, en cierto punto, de una validación, por ello, el paquete Validación está relacionado con todos aquellos paquetes que en algún momento acepten el ingreso de datos.

La **Figura 22** muestra todos paquetes existentes y las interacciones implicadas. Debido a que el sistema está diseñado para usuarios registrados, el paquete de autenticación interactúa con todos los demás.

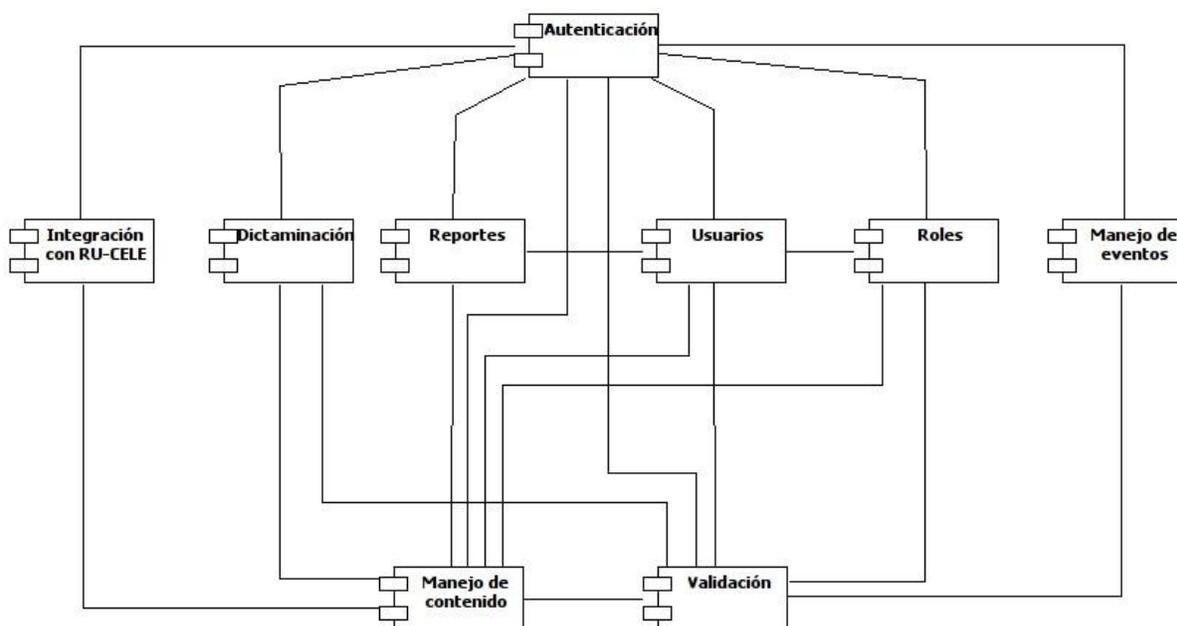


Figura 22 - Diagrama de paquetes

Para comprender mejor el diagrama de la **Figura 22** se expone, a continuación, una breve descripción de cada paquete:

- **Usuarios:** Este paquete se engloban todas aquellas acciones que necesiten obtener información acerca de los usuarios. También es el paquete encargado de realizar acciones sobre los mismos, como agregar, eliminar o modificar cierta información acerca de un usuario en específico.
- **Roles:** Los usuarios de la Didactiteca están divididos en roles, cada uno de los cuales tiene atribuidos ciertos privilegios u obligaciones. Un usuario puede tener uno o más roles, pero ciertas acciones, como el proceso de dictaminación, requieren tomar en cuenta restricciones o exclusiones que existen entre dichos roles. Este paquete se encarga de proveer información acerca de los roles de los usuarios y de las acciones que se le tienen permitido a un rol en específico.
- **Autenticación:** Para que un usuario pueda hacer pleno uso de la Didactiteca debe autenticarse en ella. El paquete de Autenticación permite a los usuarios verificar su identidad en el sitio. Pero otro aspecto importante es que permitirá el acceso a ciertas vistas dentro de la Didactiteca que solo algunos usuarios, con un rol en particular, tienen

acceso. Por ejemplo, verificará que solo los administradores puedan publicar eventos en la Didactiteca, para ello requerirá de interacción constante con el paquete de Roles.

- Manejo de contenido: El punto clave de la Didactiteca son sus contenidos, por lo cual existe un paquete dedicado únicamente al manejo de los mismos. Su extensión es amplia, puesto que la mayoría de las acciones que se realizan en la Didactiteca tendrán alguna relación con los contenidos. Dentro de este paquete se contempla la agregación y eliminación de contenido, así como el acceso a los metadatos para poder listar y buscar contenido.
- Dictaminación: Como su nombre lo indica, este paquete se encarga de todo lo referente a los procesos de dictaminación. La dictaminación se realiza sobre los contenidos, existiendo dos vertientes:
 1. Dictaminación de contenido nuevo (agregar contenido), el cual involucra la revisión de material corregido.
 2. Dictaminación de contenido existente (eliminar contenido).
- Reportes: Este paquete se encarga de realizar las acciones que impliquen el manejo y despliegue de reportes estadísticos a los Administradores. Los reportes requieren información acerca de los contenidos y de los usuarios, así que se relaciona con sus respectivos paquetes.
- Manejo de eventos: La Didactiteca notifica, de manera adicional, acerca de eventos temporales a sus usuarios (conferencias, talleres, ferias, etc.), los cuales son gestionados por los administradores. Las acciones que implica la administración de eventos serán realizadas por el paquete Manejo de eventos, desde su creación, pasando por una posible edición y finalizando con su eliminación automática al exceder la fecha del evento.
- Validación: Toda acción (caso de uso) que requiera el ingreso de datos requerirá también de una validación antes que se realice cualquier otra acción sobre los mismos. Este paquete se encargará de validar todo dato que ingrese el usuario a la Didactiteca, tal como los datos de registro de nuevo Miembro o la información de un contenido antes de ser sometido a dictaminación.
- Integración con RU-CELE: La Didactiteca será capaz de comunicarse con el repositorio institucional del CELE, llamado RU-CELE, para publicar un grupo selecto de actividades. Como tal, este paquete se encargará de dicha comunicación, tomando en cuenta las adecuaciones necesarias para hacer concordar con el esquema de metadatos de RU-CELE.

Para la descripción conceptual de los datos contenidos en la Didactiteca se ha utilizado el modelo Entidad-Relación (ER), el cual supone identificar todos los objetos, tanto físicos como lógi-

cos, que serán almacenados en la base de datos (entidades), las propiedades de interés que cada objeto posee (atributos) y la forma en que estos objetos interactúan o se vinculan entre sí (relaciones). Con ello se obtiene un modelo independiente del medio de almacenamiento y de las técnicas de acceso a los datos (27).

La **Figura 23** muestra el diagrama ER de la Didactiteca de Portugués, el cual se presenta de una manera general, debido a que la agregación de detalles, tales como atributos, generaría un diagrama más extenso que pudiera terminar siendo ilegible por sus dimensiones.

Como puede observarse, la entidad con mayores relaciones es el Contenido, puesto que gran parte de las acciones giran en torno a los recursos disponibles en la Didactiteca, además, la mayoría de las características atribuibles a un contenidos (contexto, extensión geográfica, competencia, tarea, etc.) consta de valores predefinidos para proveer una clasificación homogénea que permita realizar búsquedas de forma sencilla.

También cabe notar que en el diagrama existen dos entidades que no cuentan con una relación explícita, éstas pertenecen a los eventos temporales (conferencias, talleres, ferias, etc.) que serán mostrados en el sitio web de la Didactiteca. Solo los Administradores podrán agregar eventos, pero esta es una cuestión de seguridad y no parte de la descripción conceptual de los datos, puesto que no es necesario conocer quién publica el evento.

Este diseño asegura una implementación en donde el sistema se encuentre estrechamente vinculado tanto al plan curricular como al proceso de dictaminación para garantizar la calidad de los contenidos. Además, el diseño contempla la posible expansión del sistema, junto con su vinculación a RU-CELE, el cual es un sistema que concentra materiales didácticos de apoyo a la docencia, generados por la comunidad del CELE.

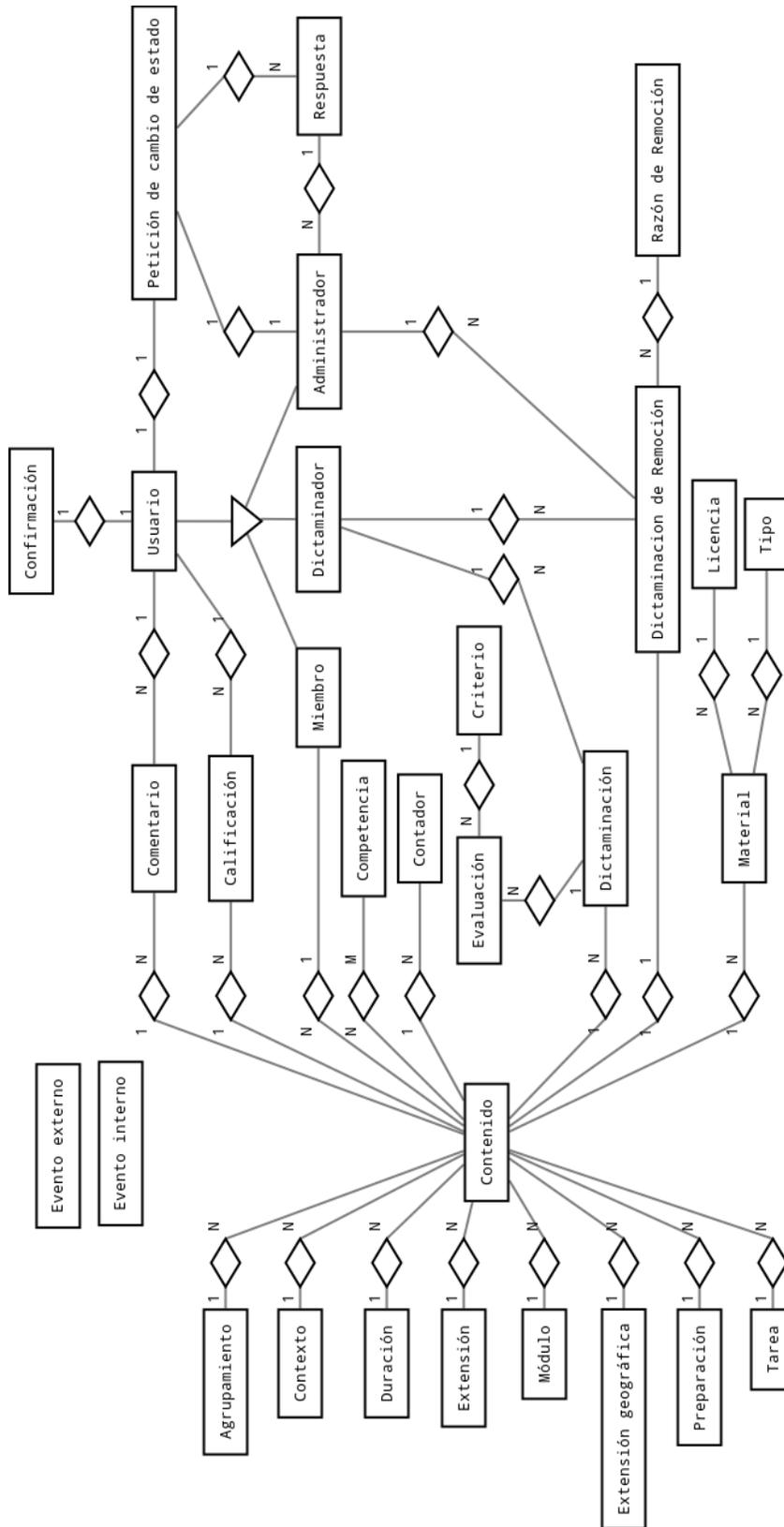


Figura 23 - Diagrama ER de la Didactiteca de Portugués

IMPLEMENTACIÓN

6.1 Arquitectura

Como arquitectura de diseño de software se utilizó el patrón Modelo Vista Controlador (MVC), el cual es un modelo de abstracción de software que considera dividir una aplicación en tres módulos según su funcionalidad: El Modelo, la Vista y el Controlador (58).

El Modelo es un conjunto de componentes que representan la información del mundo real que el sistema debe procesar. Idealmente, no debe tener relación con ninguna otra entidad dentro de la aplicación. Es decir, el Modelo desconoce la existencia de las Vista y del Controlador.

Este enfoque permite una completa independencia de los datos con la lógica de programación. Aunque, por supuesto, un elemento del sistema no puede estar completamente aislado del mismo, por ello, en la práctica el Modelo está dividido en dos submódulos (58):

- Modelo del Dominio
- Modelo de la Aplicación

El Modelo del Dominio es el conjunto de componentes representativos de la información que la aplicación va a gestionar. Mientras que el Modelo de la Aplicación proporciona los mecanismos necesarios para que el Modelo del Dominio pueda comunicarse con la Vista y el Controlador.

Las Vistas se encargan de mostrar al usuario la información del estado actual del Modelo. Una vista debe estar asociada, al menos, a un modelo, pudiendo existir varias Vistas asociadas al mismo modelo.

Una Vista obtiene del Modelo solamente la información que necesita mostrar al usuario y se actualiza conforme el Modelo del Dominio cambia, por medio de notificaciones generadas por el Modelo de la Aplicación.

El Controlador es el módulo encargado de recibir las órdenes del usuario y gestionar la aplicación en concordancia con las reglas del negocio. A partir de las órdenes que recibe, el Controlador se encarga de modificar el Modelo o de proporcionar Vistas al usuario.

La **Figura 24** muestra cómo es que se realizan las interacciones entre los distintos módulos, en donde el Controlador es el componente gestor central, puesto que la Vista, idealmente, no interactúa de manera directa con el Modelo.

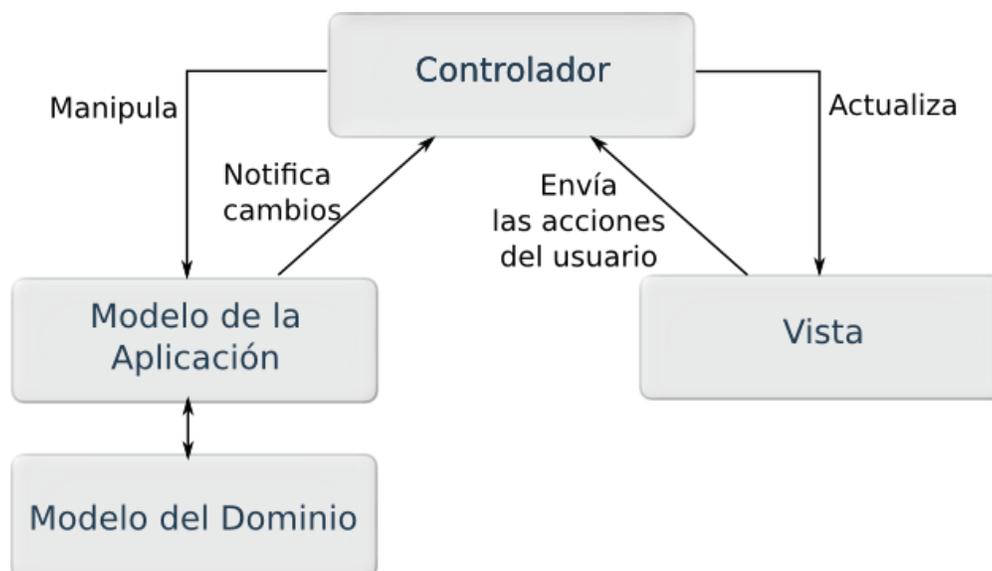


Figura 24 - Modelo MVC (58)

Para implementar el sistema de la Didactiteca se ha optado por un desarrollo propio, es decir, la Didactiteca de Portugués no estará basada en un CMS existente, puesto que requiere implementar procesos y medidas de control específicos que los sistemas de gestión actuales no poseen o cuya adaptación a éstos resultaría contraproducente, tanto en esfuerzo como en el resultado esperado.

Para ello se utilizará la plataforma de desarrollo web *Grails*^[13], la cual se distribuye libremente bajo la licencia Apache License Version 2.0.

La plataforma de Grails trabaja sobre la máquina virtual de Java y hace uso de otras plataformas de las cuales se conoce de antemano su versatilidad y buen funcionamiento, debido a que han sido extensamente probadas y utilizadas en otros entornos (59). Tales plataformas son:

- Spring: Para el flujo de trabajo e inyección de dependencias en tiempo de ejecución.
- Hibernate: Para la persistencia de los datos y mapeo de objetos a una base de datos relacional.
- SiteMesh: Para la composición de las vistas con un esquema consistente.

Grails provee, además, varias características que son de especial interés para desarrolladores que requieran un entorno productivo, consistente y de relativa facilidad, tanto en su comprensión como uso (60). Las características que permiten dicho entorno son:

- Paradigma de convención sobre configuración: Grails trabaja sobre un conjunto de convenciones bien definidas que pretenden disminuir el número de parámetros que un desarrollador debe configurar.

¹³ <http://grails.org/>

- Documentación relevante: El sitio oficial de Grails provee una documentación de todos los métodos que un desarrollador requerirá, junto con ejemplo de uso.
- Soporte para Ajax^[14] extensible y personalizable.
- Entorno completo de desarrollo: Grails incluye un servidor web y una base de datos de prueba que son generados y cargados automáticamente desde la primera ejecución del sistema. Tales características no requieren de configuración alguna. Es decir, un desarrollador puede crear una versión preliminar del sistema y ponerla a prueba con la sola ejecución de un comando.
- Métodos dinámicos: Es posible hacer uso de métodos que son generados por en tiempo de ejecución, lo cual remueve del desarrollador la carga de escribir código para tareas comunes o repetitivas, como, por ejemplo, las funciones básicas relacionadas con la gestión de los datos persistentes.

Aunados a estas características están los modelos y paradigmas utilizados por Grails. Al realizar el diseño de la aplicación, se ha utilizado el estándar UML, descrito en el capítulo 5.3, el cual está enfocado en sistemas descritos bajo el paradigma de la programación orientada a objetos. Grails, al trabajar sobre la máquina virtual de Java, permite un desarrollo orientado a objetos mediante el lenguaje de programación Groovy^[15]. Con ello, la implementación lógica puede ser fácilmente obtenida de los diseños UML.

Además, Grails utiliza el patrón MVC de un modo que asemeja casi en su totalidad al flujo descrito en la **Figura 24**, pero con la añadidura de un nuevo elemento: los Servicios. Un servicio engloba, dentro de un mismo objeto, los procesos relacionados con las reglas del negocio que requieran ser utilizados en más de un contexto, es decir, son aquellos procesos que requieran modificar el Modelo de manera extensa o que necesiten ser implementados en diversos Controladores (61).

6.1.1 Modelo de datos

Partiendo del modelo ER expuesto en la **Figura 23** del capítulo 5.3, se generó el modelo relacional, cuya representación gráfica se encuentra dividida entre las **Figura 25** y **Figura 26**, debido a la gran cantidad de relaciones y atributos existentes.

Dado que se hará uso de la plataforma de desarrollo Grails para la implementación del sistema, fue necesario contemplar diversas convenciones y restricciones propias de esta plataforma.

¹⁴ AJAX no cuenta con una página oficial, pero una recopilación de varias páginas relacionadas al tema puede observarse en: <http://www.dmoz.org/Computers/Programming/Languages/JavaScript/Ajax/>. Otra fuente importante de recursos es el sitio de ASP.NET: <http://www.asp.net/ajax>

¹⁵ <http://groovy.codehaus.org/>



Figura 25 - Modelo relacional

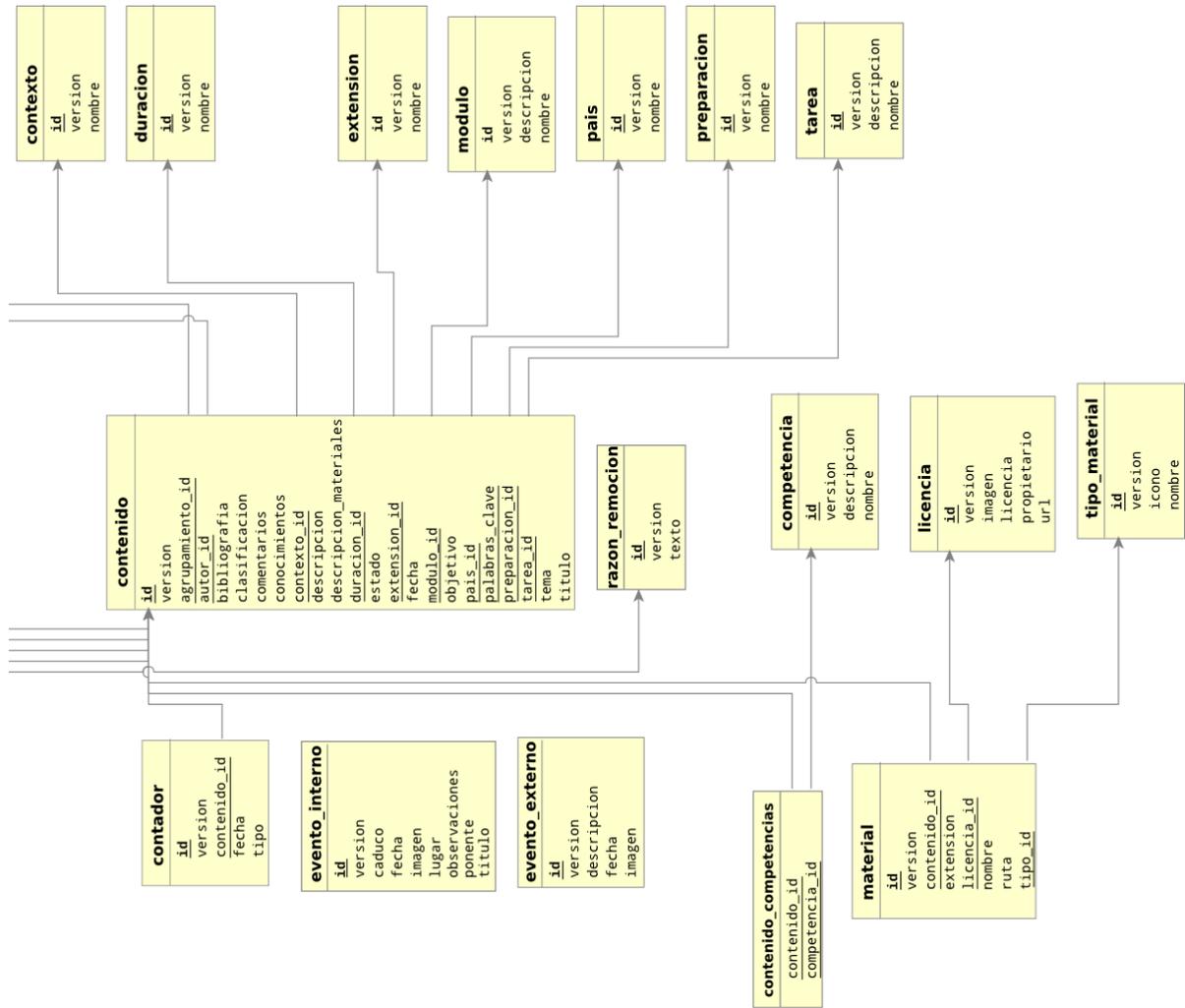


Figura 26 - Modelo relacional

Existen tres aspectos fundamentales a considerar cuando se hace uso de Grails (60):

- El modelo del dominio está configurado para utilizar claves primarias de valor numérico entero.
- No se recomienda el uso de llaves compuestas, ya que será responsabilidad del desarrollador implementar todas las funciones necesarias para la interacción entre la plataforma de desarrollo y el DBMS que requieran el uso de llaves.
- Para soportar correctamente peticiones concurrentes, se agrega el atributo *version*, cuyo valor inicial comienza en 0 y se incrementa con cada modificación que se realice a una tupla.

Tomando en cuenta los puntos anteriores se ajustó el modelo relacional para contemplar aquellos atributos necesarios para una total integración con el modelo del dominio estándar de Grails: *id* y *version*.

El esquema relacional ha sido normalizado hasta la forma normal Boyce-Codd (27), en donde cabe destacar que el cumplimiento de dicha forma normal ha sido facilitado debido al uso de las convenciones que maneja el modelo del dominio en Grails, en donde se busca evitar la presencia de llaves compuestas.

6.1.2 Despliegue de la Didactiteca de Portugués

La Didactiteca está implementada en un servidor *Jakarta Tomcat* (62), el cual contiene y ejecuta todas las instrucciones del sistema. Tomcat se conecta con un servidor *Apache HTTP* (63), el cual permite un control más granular de las transacciones web y gestiona las conexiones con el sistema.

La conexión entre ambos servidores se realiza mediante un módulo especial, *mod_jk*, que permitirá la interacción entre los servidores, así como el paso de peticiones del usuario hacia la Didactiteca y el envío de su respectiva respuesta, puesto que es el servidor Apache es quien se comunica con el navegador web que reside en el equipo de un usuario, para poder permitir el acceso a los servicios.

Para almacenar los datos de usuarios, contenidos, roles, eventos, etc. la Didactiteca utilizará una base de datos relacional cuyo DBMS es *PostgreSQL* (64), para el cual se cuenta actualmente con un entorno apropiado para el desarrollo y puesta en marcha, gracias a proyectos anteriormente desarrollados en el CAD, por lo cual la integración de la base de datos con el sistema de la Didactiteca se realizó sin mayores complicaciones.

Todos los archivos tanto aquellos propios de la aplicación, como los materiales que componen a las actividades y las imágenes que acompañan a los eventos, residirán en el disco duro del servidor web que contiene al sistema y su acceso estará permitido únicamente a usuarios registrados.

La **Figura 27** muestra el diagrama de despliegue de la Didactiteca de Portugués, en donde se especifican las interacciones existentes entre los distintos componentes que conforman al servidor.

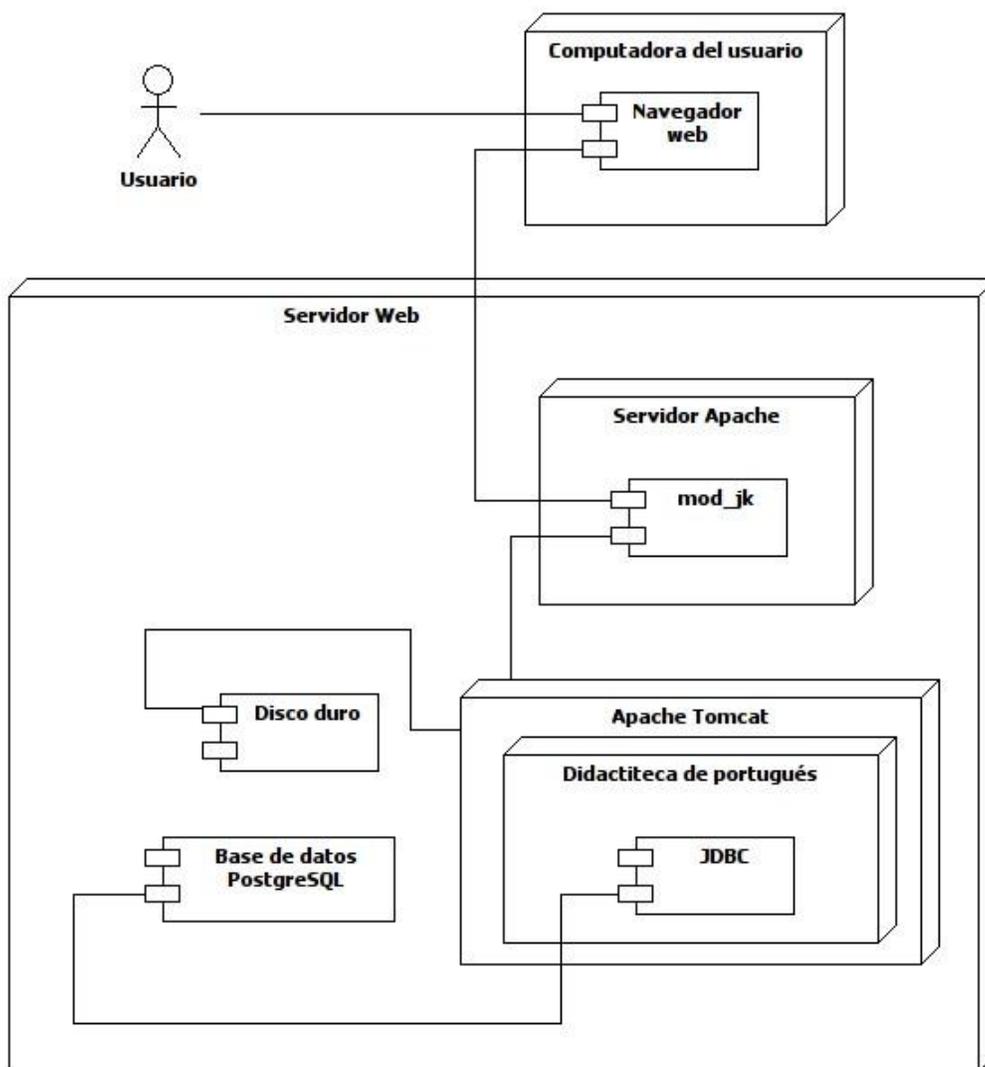


Figura 27 - Diagrama de despliegue de la Didactiteca de Portugués

Se ha mencionado que Grails es una plataforma libre, lo cual implica que es distribuida sin costo alguno y puede utilizarse tanto en medios educativos como comerciales, siempre y cuando se respeten los términos de la licencia. Este mismo concepto aplica para todas las demás plataformas, herramientas y elementos utilizados en el sistema de la Didactiteca de Portugués, los cuales han sido escogidas teniendo siempre en mente su tipo de licencia, para garantizar que el sistema pueda ser implementado sin que exista la posibilidad de incurrir en una falta.

El uso de software libre ha sido dispuesto como un requisito, ya que permite elaborar un sistema desapegado de cualquier pago por licenciamiento o derechos de autor.

6.2 Desarrollo técnico

Las Vistas contenidas en la Didactiteca de Portugués serán presentadas a los usuarios como páginas web, por lo tanto se ha tenido el cuidado de seguir diversas técnicas para la implementación de las Vistas, de modo tal que se conserve íntegra la funcionalidad del sistema pero sin comprometer de modo alguno el diseño visual o el comportamiento de los elementos que componen dichas Vistas.

Recientemente, el entorno web ha adoptado especificaciones como *Hyper Text Markup Language Revision 5 (HTML5)* y *Cascading Style Sheets Level 3 (CSS3)*, tratándolos como estándares a los cuales todo nuevo desarrollo web debería apegarse. Sin embargo, ninguno de ellos, al tiempo en que se redacta esta tesis, es un estándar (65).

HTML5 está clasificado por el World Wide Web Consortium (W3C), la organización emisora, como un trabajo en progreso, cuyo desarrollo ha sido continuo durante los últimos años y del cual se esperan únicamente cambios menores. Actualmente puede verse presente en, aproximadamente, 74% del mercado competente a los exploradores web (66). Debido al gran uso e impacto de esta tecnología, la Didactiteca ha sido creada para apegarse a las especificaciones actuales de HTML5, pero sin mellar la funcionalidad web de aquellos usuarios que utilicen exploradores antiguos o que no implementen las funcionalidades esperadas, por ejemplo, Internet Explorer 8 o inferior.

En cuanto a la especificación CSS3, existe una disparidad impresionante, puesto que algunos componentes están clasificados, por la W3C, como estándares, mientras que otros siguen siendo borradores (65). Todo ello ha resultado en una enorme diversidad de implementaciones en los navegadores actuales, puesto que cada uno soporta distintos componentes cuya nomenclatura, en ocasiones, difiere de la sugerida por la especificación CSS3. Debido a esto, la Didactiteca, en lo competente al apartado visual y presentación de las vistas, posee un enfoque centrado en la compatibilidad. Es decir, las Vistas se apegarán lo más que sea posible a la especificación CSS3, pero cuidando que la presentación y disposición de los elementos sean homogéneos entre los exploradores más utilizados actualmente.

La Didactiteca ha sido diseñada para funcionar correctamente con los exploradores que implementen parcial o totalmente las especificaciones de HTML5 y CSS3:

- Cualquier explorador basado en Webkit versión 534 o superior (67).
 - Verificado con Chrome 18, Chromium 18 y Safari 5.1
- Firefox 3.6 o superior (68).
 - Verificado con Firefox 14 y 15

- Internet Explorer 8 o superior. Aunque la versión 8 de Internet Explorer no soporta de ninguna manera HTML5 o CSS3, es un explorador que cuenta aproximadamente con el 22% del mercado (66). Es la última versión que acepta el sistema operativo Windows XP, y se incluye específicamente por cuestiones de compatibilidad.
 - Verificado con IE 8 y 9

En total, los exploradores utilizados para la verificación comprenden entre el 88% y 95% de la presencia en el mercado a nivel mundial, según exponen diversos reportes estadísticos (66) (69) (70), aunque es difícil conocer con exactitud el estado actual del mercado en los exploradores web, es posible asegurar un margen óptimo de compatibilidad al diseñar un sistema capaz de funcionar adecuadamente al ser accedido desde los navegadores web más utilizados.

Otra característica de relevancia, que se relaciona tanto con el aspecto visual y la presentación de Vistas como el almacenamiento de información, es el soporte de símbolos diacríticos, los cuales corresponden a los acentos en el español, aunque existen otros símbolos como la cedilla (ç) que son utilizados en el portugués. Debido a que la Didactiteca implementa, al menos, dos idiomas, portugués y español, es crucial que toda la información pueda ser desplegada correctamente a los usuarios. Por ello, la Didactiteca utiliza internamente el formato de codificación *8bit Unicode Transformation Format (UTF-8)* para la representación de caracteres. UTF-8 utiliza el estándar de caracteres *Unicode* que incluye, entre otros, los caracteres y símbolos diacríticos del alfabeto latino, utilizado por las lenguas romances como el español y portugués (71).

6.2.1 Integración con RU-CELE

En el capítulo 5.3, donde se listan todos los casos de uso de la Didactiteca, puede observarse que la integración con RU-CELE es uno de ellos, pero se ha decidido descartarlo ya que los objetivos de ambos proyectos difieren en cuanto al público objetivo.

RU-CELE está diseñado para centralizar materiales generados únicamente por la comunidad del CELE (72), mientras que la Didactiteca de Portugués tiene considerada una posible expansión a entidades externas de enseñanza del idioma portugués.

En primera instancia la Didactiteca estará a disposición únicamente de los profesores del departamento de portugués en el CELE. En esta fase pudiera existir la integración con el repositorio RU-CELE. Pero es imposible dejar de lado la expansión de la Didactiteca, que generaría un conflicto tanto con el CELE como con el autor de un material. En tal caso, RU-CELE estaría dando difusión a un material ajeno a la institución, mientras que la Didactiteca estaría permitiendo la atribución de los méritos a la entidad errónea, es decir, un material estaría públicamente atribuido al CELE, aun cuando la persona no tenga relación alguna con este centro.

Además, la Didactiteca de Portugués es un sistema predominantemente cerrado, cuyos materiales están disponibles, en su totalidad, a usuarios registrados; el público únicamente tiene acce-

so a vistas previas, junto con la posibilidad de registrarse. Mientras que RU-CELE es un repositorio público donde cualquier persona puede conocer el material intelectual que ahí se aloja. Permitir que una porción del material de la Didactiteca sea completamente público crearía una excepción, lo cual no es deseable, puesto que genera un flujo alternativo que se debe analizar cuidadosamente para evitar cualquier choque con las reglas de negocio actuales.

Por tal motivo, cualquier material que algún profesor desee hacer público deberá ser publicado en RU-CELE directamente, sin intervención o mediación alguna por parte de la Didactiteca de Portugués.

6.2.2 Pruebas de rendimiento

Terminadas las fases de diseño e implementación del sistema, la Didactiteca se sometió a un constante proceso de pruebas para determinar el rendimiento del sistema y realizar optimizaciones en los procesos que pudieran estar demorando demasiado tiempo en ejecutarse. Tal proceso es conocido como *pruebas de rendimiento*.

Una prueba de rendimiento consiste en evaluar la respuesta de un sistema bajo ciertas condiciones de operación, usualmente con la ayuda de alguna herramienta de software que permita automatizar tales condiciones (73).

Se considera que el rendimiento de un sistema es una cuestión de percepción, la cual depende tanto del entorno del sistema como de sus usuarios (74). Esta denotación, ciertamente, parece no ser suficiente, puesto que toda aplicación debe ser probada antes de estar disponible y, por tanto, requiere de una base sólida sobre la cual construir las pruebas.

Sin embargo, según diversos autores, el tema de pruebas dentro del entorno de software sigue siendo un campo informal, en donde los resultados esperados dependen de muchos factores inherentes de la propia aplicación, difícilmente mesurables de un modo global y homogéneo (74) (73) (75). Por ejemplo, una de las medidas de mayor importancia para las aplicaciones web es el tiempo que tarda el servidor en responder a una petición. Podría decirse que si una aplicación web responde en un tiempo menor o igual a n segundos, se le considera de buen rendimiento, pero esta cantidad podría decrementar con el paso del tiempo, debido a redes más veloces y equipos más potentes. También existen casos donde un usuario está dispuesto a esperar más tiempo, por ejemplo, al concretar una compra en línea, el servidor debe verificar y realizar cargos a una tarjeta de crédito, actualizar sus registros e iniciar el proceso de envío; el usuario ya no requiere realizar acción alguna, así que no tiene mayor problema si la aplicación tarda unos segundos más en comparación con el resto del proceso de compra (75).

Todo ello dificulta el determinar un conjunto de valores a los cuales apegarse. Por tanto no es posible asignar al rendimiento una magnitud fija. Sin embargo, es posible medir y calcular el rendimiento de una aplicación, la consideración a tomar es no tratar cumplir al pie de la letra

alguna regla cuyo fundamento no esté relacionado directamente con la aplicación (por ejemplo, cifras calculadas en alguna publicación, proyectos ajenos con poca o ninguna similitud, etc.).

«Una aplicación con un buen rendimiento es aquella permite al usuario final realizar una determinada tarea sin dilaciones indebidas o frustración.» (74)

Si bien es cierto que el rendimiento depende en gran parte de la perspectiva del usuario final, es necesario considerar una serie de indicadores para poder medir el rendimiento de un sistema. Para las aplicaciones web, son cuatro los indicadores que suelen vigilarse con detenimiento para garantizar el buen rendimiento del sistema: (73) (74)

- Disponibilidad: La cantidad de tiempo que una aplicación está disponible para el usuario final. La falta de disponibilidad es importante porque muchas aplicaciones tendrán un costo sustancial para el negocio, incluso un corte pequeño. En cuanto a las pruebas de rendimiento, esto significaría la total incapacidad de un usuario para hacer un uso efectivo de la aplicación.
- Tiempo de respuesta: La cantidad de tiempo que tarda la aplicación para responder a una petición de usuario. Para las pruebas de rendimiento, normalmente se mide el tiempo de respuesta del sistema, que es el tiempo entre la petición del usuario y una respuesta que llega a la estación de trabajo del usuario.
- Tasa de trabajo: La velocidad a la que un evento ocurre. Todo evento debe tener relación con un usuario. Por ejemplo, el número de visitas a una página web dentro de un período de tiempo determinado.
- Utilización: El porcentaje de capacidad teórica de un recurso que está siendo utilizado. Por ejemplo, la cantidad de ancho de banda que consume una aplicación y la cantidad de memoria utilizada en un servidor con un cierto número de usuarios activos.

Anteriormente han existido propuestas e investigaciones al respecto de cada uno de estos indicadores. Por ejemplo, desde la perspectiva del usuario final, uno de los aspectos más relevantes es el tiempo de respuesta.

A lo largo de los años se han llevado a cabo diversas investigaciones para calcular o sugerir algún tiempo de respuesta límite. Uno de los primeros reportes que se tiene al respecto, escrito por Robert Miller, describió tres umbrales de la atención humana en lo referente a la interacción con una computadora: (76)

1. Un tiempo de respuesta de un décimo de un segundo es suficiente para considerarse instantáneo.
2. Una respuesta dentro de 1 segundo es lo suficientemente rápida como para que los usuarios sientan que están interactuando libremente con la información.
3. Los tiempos de respuesta deben mantenerse por debajo de 10 segundos para mantener la atención del usuario.

Mientras tanto, para el caso específico de las páginas web, se mantuvo, durante muchos años un límite superior ideal de 8 segundos, de acuerdo a las pruebas realizadas por Peter Bickford (77). Pero este límite fue reducido a la mitad, en el año 2006, por una nueva investigación por parte de las empresas Akamai y JupiterResearch, las cuales enfocaron su estudio en sitios de compras en línea (78).

A pesar de que estas cifras han tenido trascendencia a lo largo de los años, es la experiencia del usuario la que define un límite apropiado. Por ello se retomaron los proyectos analizados en el capítulo 4.3 para realizar una comparativa del tiempo que un usuario podría esperar en cuanto al rendimiento de un sitio web enfocado a proveer recursos educativos.

Para la realización de este análisis, se utilizaron las herramientas de desarrollador que los exploradores Chrome y Chromium poseen^[16], en donde es posible conocer el tiempo de respuesta de una página web ante una petición. Los resultados pueden observarse en la **Tabla 8**. Durante las pruebas se ha deshabilitado el uso de cache en el explorador, para evitar falsas mediciones por contenido previamente almacenado en el equipo.

	Página inicial [s]	Búsqueda rápida [s]	Búsqueda avanzada [s]	Visualizar contenido [s]
MIT OCW	3.05	1.51	2.40	2.90
MERLOT	3.96	7.8	10.6	2.74
CREA	3.46	7.20	7.11	4.1
DAR	2.16	2.25	2.10	2.24
Funes	3.6	13.15	8.91	5.04
Promedio	3.24	6.38	6.22	3.39

Tabla 8 - Tiempos de respuesta de acervos y repositorios

¹⁶ La documentación de estas herramientas se encuentra disponible en el sitio oficial del proyecto Chromium: <http://www.chromium.org/devtools>

Gracias a este análisis es posible observar que los tiempos de respuesta tienen cierta soltura en cuanto a búsquedas se trata, mientras que el simple despliegue de información no debería tomar más de 4 segundos.

Los demás aspectos de rendimientos son difíciles de obtener para estos sitios, puesto que no se tiene ningún control sobre los mismos. Sin embargo, tanto Funes como el MIT OCW ofrecen un reporte estadístico de sus sitios, de donde es posible calcular un estimado para el tasa de trabajo, medido en transacciones por unidad de tiempo, donde se considera como una transacción a cada petición que un usuario realiza al sitio.

- MIT OCW: 423,269 transacciones por mes, aprox. 9.79 transacciones por minuto (79).
- Funes: 10,000 transacciones por mes, aprox. 0.23 transacciones por minuto (52).

Estos datos consideran que las transacciones están distribuidas uniformemente a lo largo del tiempo. Es la mejor aproximación que puede obtenerse con la información disponible y provee un indicio del promedio de transacciones que una aplicación debe ser capaz de manejar.

Para probar y medir el rendimiento la Didactiteca de Portugués, el primer paso fue analizar el uso que la Didactiteca tendrá durante su primera fase, es decir, mientras sea únicamente utilizado por los profesores del Departamento de Portugués. Dicho departamento cuenta, actualmente, con 16 profesores. Por lo tanto se realizarán las pruebas iniciales suponiendo que la totalidad del profesorado está haciendo uso del sistema; aunque este caso es poco probable (tal como será analizado posteriormente) no debe descartarse y para una cantidad baja, como lo es 16 usuarios, no se generará una carga desmesurada en el servidor.

Al realizar las pruebas se ha utilizado la herramienta de software JMeter^[17], la cual permite simular peticiones a un servidor de modo tal que asemeje el flujo normal de trabajo de un usuario, mientras recaba datos estadísticos para determinar el rendimiento de una aplicación web.

La **Tabla 9** muestra el resultado de simular 16 usuarios registrados, realizando tareas comunes, entre cada petición existe un lapso de espera cuyo rango comprende de 2 a 10 segundos, el cual simula el tiempo que un usuario tarda en tomar una decisión para realizar la siguiente acción.

¹⁷ La descripción completa de esta herramienta, así como su documentación, puede obtenerse de la página oficial del proyecto JMeter: <http://jmeter.apache.org/>

Petición	Transacciones	Promedio [ms]	Mínimo [ms]	Máximo [ms]	Tasa de trabajo [trans/min] *
Iniciar sesión	16	92	74	123	37.41
Página inicial	142	64	39	93	14.70
Búsqueda avanzada	36	50	34	65	3.84
Búsqueda simple	56	122	46	168	6.10
Listar contenido	44	151	121	198	4.80
Ver contenido	132	95	64	168	14.40
Comentar contenido	130	106	65	142	14.69
TOTAL	556	94	34	198	55.66

* trans: transacción

Tabla 9 - Prueba de rendimiento: 16 usuarios

La **Figura 28** muestra los tiempos de respuesta agrupados por el tipo de petición, mientras que la **Figura 29** engloba todas las peticiones para denotar el tiempo promedio de respuesta. Ambas gráficas corresponden a la prueba realizada con 16 usuarios simultáneos y fueron generadas mediante la aplicación JMeter.

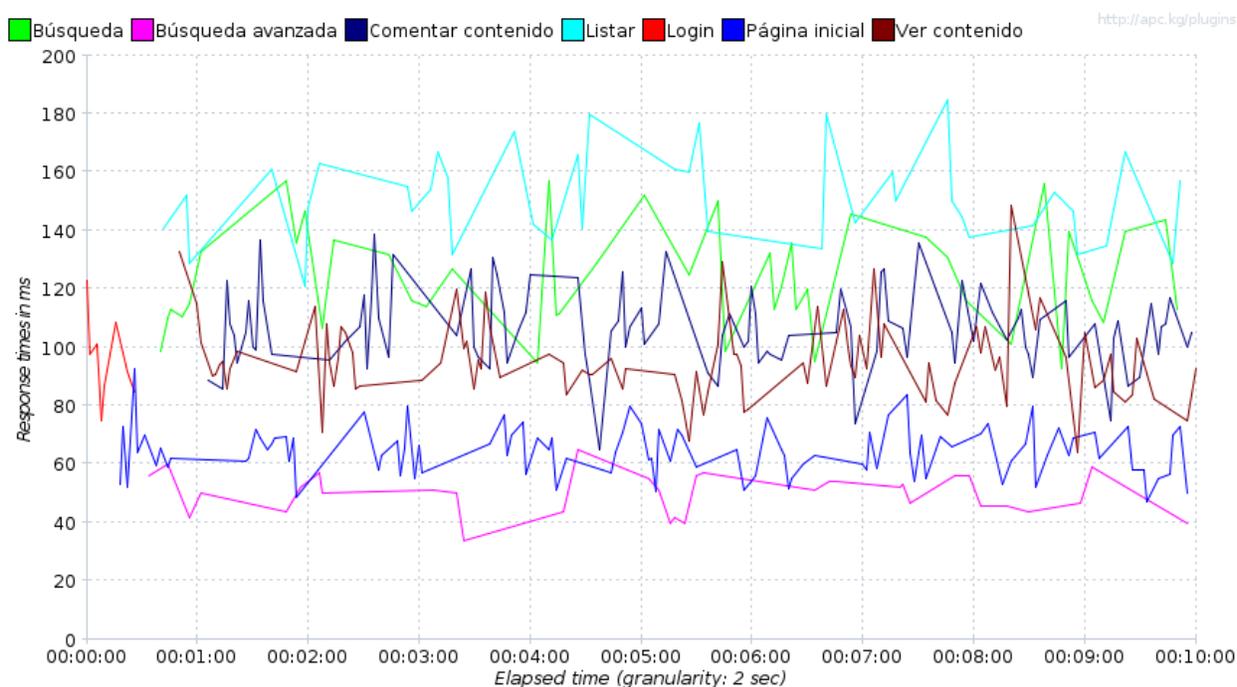


Figura 28 - Tiempos de respuesta: 16 usuarios

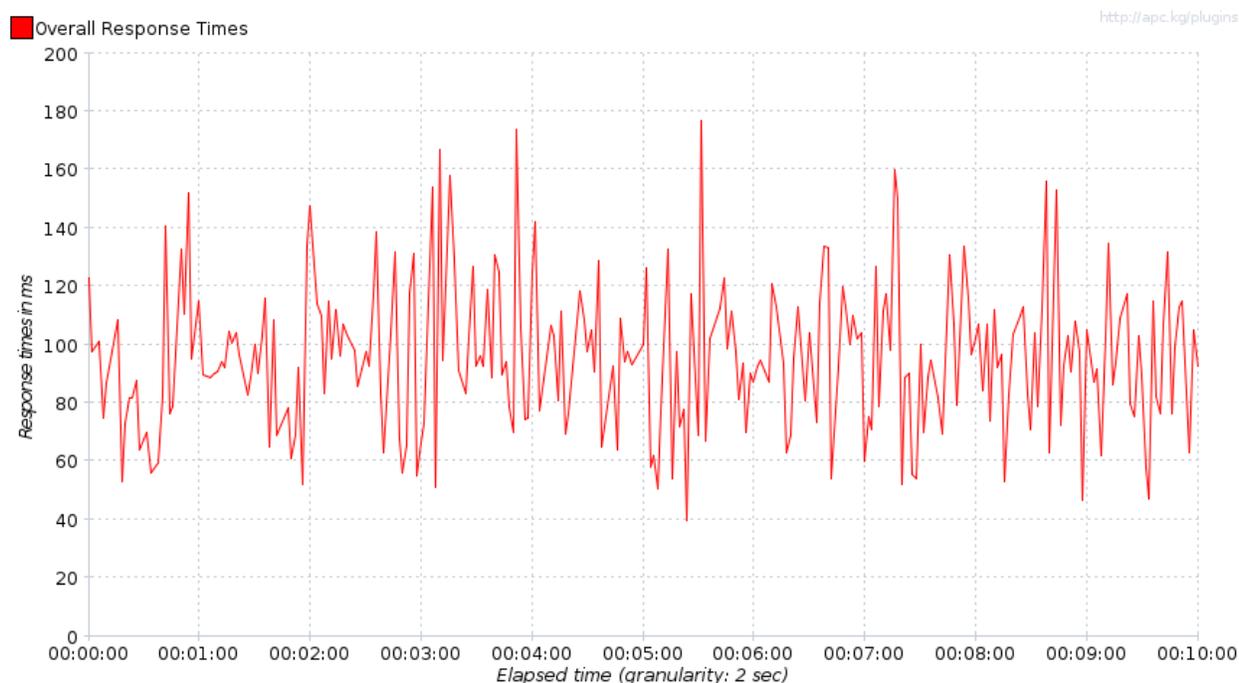


Figura 29 - Tiempo de respuesta promedio: 16 usuarios

Con 16 usuarios la Didactiteca es capaz de responder en un tiempo menor al segundo. Aunque cabe mencionar que para estas pruebas la latencia^[18] de red es mínima puesto que se trabaja bajo una red local. Pero también se realizaron pruebas desde una red externa, obteniendo siempre una diferencia promedio menor a los 200 ms con respecto a los tiempos mostrados en la **Tabla 9**.

Como segunda fase se analizó el rendimiento de la Didactiteca suponiendo una expansión hacia sedes externas, lo cual implicaría un crecimiento en cuanto a los usuarios registrados en el sistema. Tomando como ejemplo uno de los proyectos más ampliamente difundidos y del cual se cuenta con mayor información estadística, el MIT OCW, sabemos que, en promedio, cada usuario realiza 9.1 transacciones durante todo el tiempo que permanece en el sitio de MIT OCW; además, el tiempo promedio de permanencia equivale a 9 minutos. Si consideramos una distribución uniforme, podemos hacer uso del tasa de trabajo mencionado anteriormente (9.79 transacciones por minuto) para calcular que, en promedio, existen 9.68 usuarios presentes en el sitio de MIT OCW en un minuto dado, cada uno de ellos realizando 1 transacción por minuto.

Como una segunda referencia se encuentra el proyecto de Funes, cuyo tasa de trabajo es de 0.23 transacciones por minuto, una cantidad mucho menor que el caso del MIT OCW. Sin embargo, Funes registra un aproximado de 1,000 usuarios en su reporte estadístico (52) (el MIT OCW no cuenta con un sistema de registro para poder realizar la comparación). Si, además, tomáramos como referencia el comportamiento de los usuarios del MIT OCW para generar una aproximación de uso en Funes, se obtendría un promedio de 0.22 usuarios simultáneos para un

¹⁸ Latencia: medida de tiempo que indica el retardo total en una red.

minuto dado. Esta cantidad indica que durante el transcurso de un día 327.56 usuarios podrían haber utilizado el sistema (máximo teórico).

Con ello podemos concluir que la cantidad de usuarios que hacen uso simultáneo de un sistema, inclusive uno tan grande como el MIT OCW, es menor de lo que inicialmente pudiera imaginarse dada la enorme cantidad de visitas que recibe el sitio. Este hecho suele ser el origen de muchos errores al momento de realizar pruebas de rendimiento, en donde inicialmente podría considerarse que si una aplicación tiene n usuarios, entonces debe ser probada para soportar al menos n usuarios simultáneos (80).

Se espera que la Didactiteca aumente su alcance para poder abarcar más instituciones, el período en que esto suceda es desconocido, también lo es la cantidad de usuarios que resulten de tal crecimiento. Sin embargo, es posible fijar una meta a mediano plazo equivalente a la cantidad de usuarios de Funes.

Por ello, se consideró un escenario donde se cuente con 1,000 usuarios registrados, de los cuales, apoyándonos con los cálculos realizados para las estadísticas de uso de Funes, se considerará que el 30% de ellos estarán accediendo al sistema de manera simultánea y continua. Las acciones y lapsos de espera serán los mismos que en la primera prueba. Este escenario supone un uso mayor del sistema que bajo condiciones normales, aunque sin ser desmesurado; ya que es de gran utilidad para asegurar que la aplicación seguirá funcionando aun en momentos *pico*, es decir, cuando surja un súbito incremento de peticiones al sistema.

Los resultados de la segunda prueba pueden observarse en la **Tabla 10**.

Petición	Transacciones	Promedio [ms]	Mínimo [ms]	Máximo [ms]	Tasa de trabajo [trans/min]
Iniciar sesión	300	84	57	356	120.49
Página inicial	1,939	55	34	392	199.61
Listar	612	161	99	591	66.10
Búsqueda avanzada	598	38	24	190	63.80
Búsqueda	664	121	71	671	71.66
Ver contenido	1,790	105	60	688	198.85
Comentar contenido	1,707	123	69	610	197.43
TOTAL	7,610	96	24	688	758.16

Tabla 10 - Prueba de rendimiento: 300 usuarios

La **Figura 30** muestra los tiempos de respuesta agrupados por el tipo de petición, mientras que la **Figura 31** engloba todas las peticiones para denotar el tiempo promedio de respuesta. Ambas gráficas corresponden a la prueba realizada con 300 usuarios simultáneos y fueron generadas mediante la aplicación JMeter.

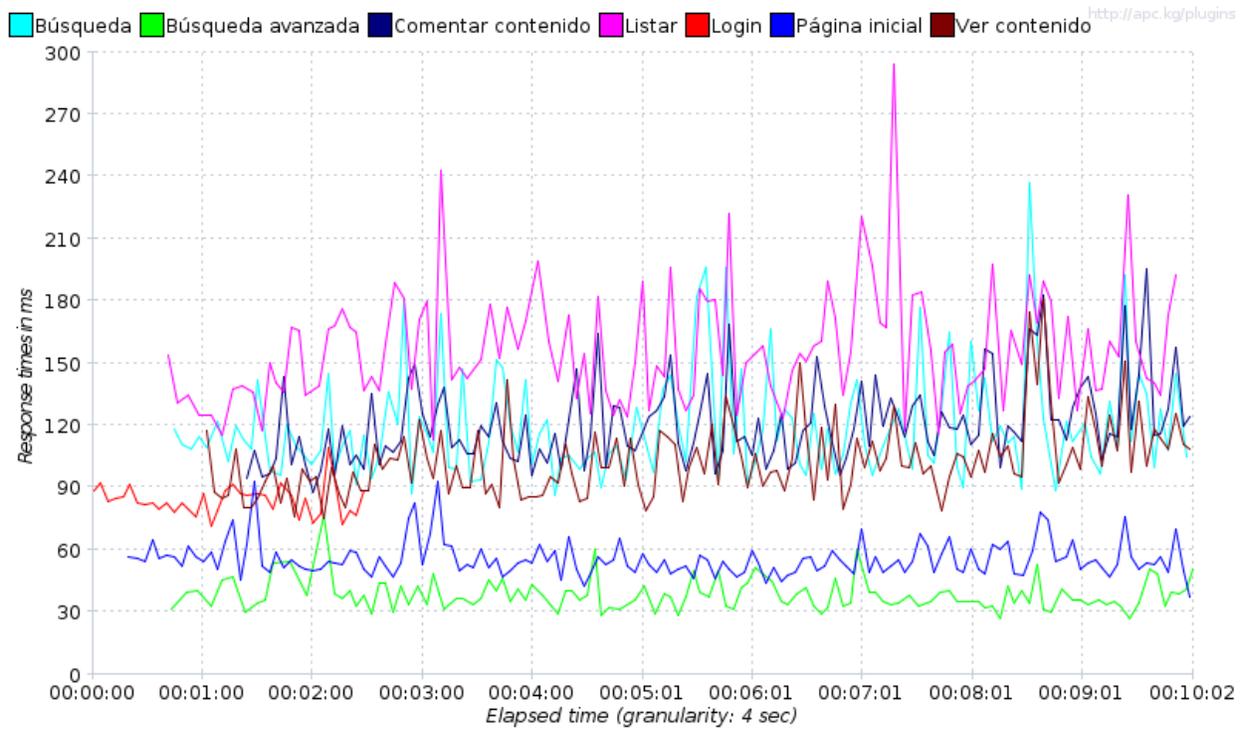


Figura 30 - Tiempos de respuesta: 300 usuarios

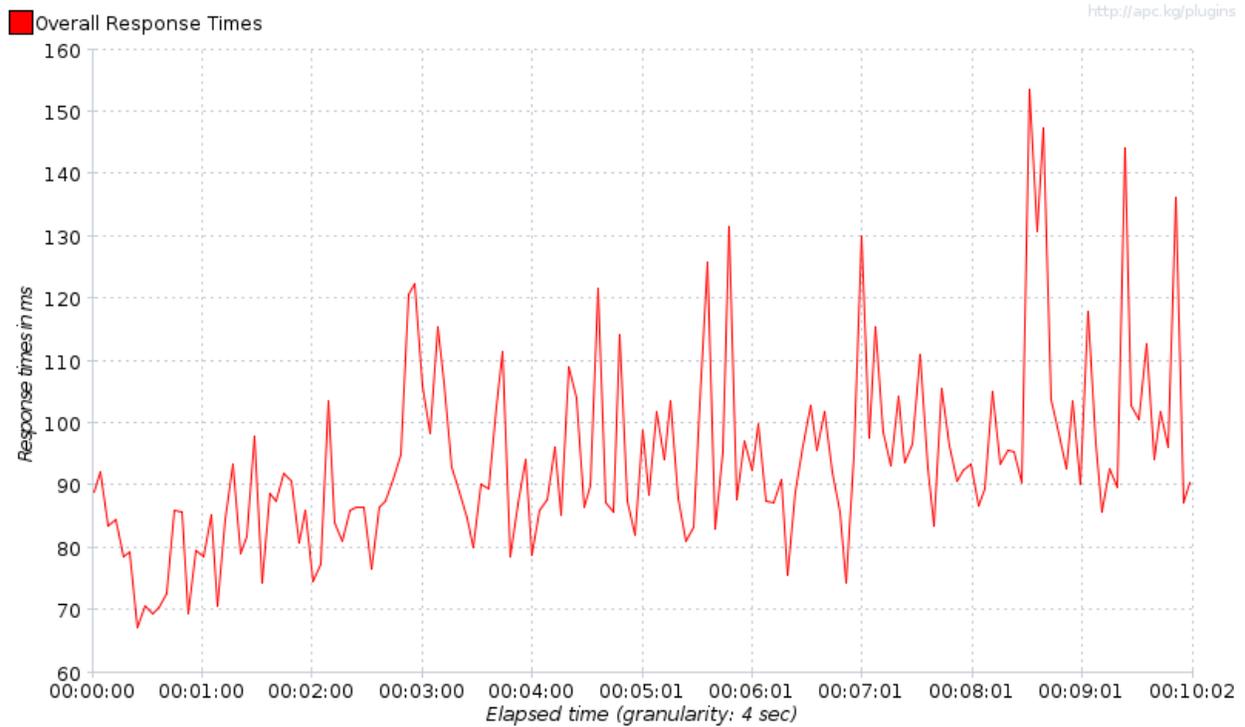


Figura 31 - Tiempo de respuesta promedio: 300 usuarios

Realizando una comparación con la prueba inicial que consideraba tan solo 16 usuarios simultáneos, es posible notar que el incremento más significativo se da en los tiempos máximos de respuesta, ya que, en promedio, el sistema se comporta de un modo similar tanto para 16 como

para 300 usuarios simultáneos, lo cual asegura un buen rendimiento de la Didactiteca tanto para la primera fase como para una expansión considerable.

Para conocer la capacidad máxima del sistema, existe una prueba de rendimiento llamada *prueba de estrés*, la cual se centra en determinar el rendimiento del sistema cuando se somete a condiciones más allá de las previstas durante la fase de desarrollo. Esta prueba está diseñada para determinar en qué condiciones una aplicación fallará, cómo va a fallar y qué indicadores pueden ser monitoreados para advertir dicha falla (73).

Para la prueba de estrés se contempló un máximo de 3,000 usuarios simultáneos, lo cual representaría aproximadamente 15,000 usuarios registrados en la Didactiteca, de acuerdo con los datos generados para la segunda prueba de rendimiento. Paulatinamente se incrementará la cantidad de usuarios que se conectan al sistema y la prueba terminará cuando la Didactiteca no sea capaz de responder a una solicitud, aunque no se alcance el límite de 3,000 usuarios.

La **Tabla 11** muestra los resultados para la prueba de estrés. Los tiempos máximos corresponden a la etapa en donde el sistema se encontraba a punto de fallar y los valores son ya inaceptables puesto que las transacciones toman más de un minuto en concretarse. Otro dato significativo es la tasa de trabajo total, donde se puede observar un valor muy cercano al límite práctico de transacciones que el sistema podría procesar en un minuto.

Petición	Transacciones	Promedio [ms]	Mínimo [ms]	Máximo [ms]	Tasa de trabajo [trans/min]
Iniciar sesión	1,174	7,707	51	77,672	58.88
Página inicial	5,423	5,693	35	63,369	278.50
Listar	1,771	6,224	99	63,431	93.23
Búsqueda avanzada	1,677	1,849	25	63,228	88.82
Búsqueda	1,730	2,361	74	63,214	92.07
Ver contenido	4,922	6,080	64	63,157	265.04
Comentar contenido	4,660	11,283	73	77,274	256.36
TOTAL	21,357	6,585	25	77,672	1,070.18

Tabla 11 - Prueba de estrés: resultados

La **Figura 32** muestra los tiempos de respuesta agrupados por el tipo de petición, debido a la gran cantidad de transacciones que la gráfica resume tal vez no sea posible ver con claridad el comportamiento de la aplicación. Sin embargo, se puede observar que existe un punto donde la aplicación comienza a demorar más sin lograr estabilizarse, esto sucede al momento que la Didactiteca cuenta con 640 usuarios simultáneos, pero los tiempos de respuesta siguen siendo tolerables.

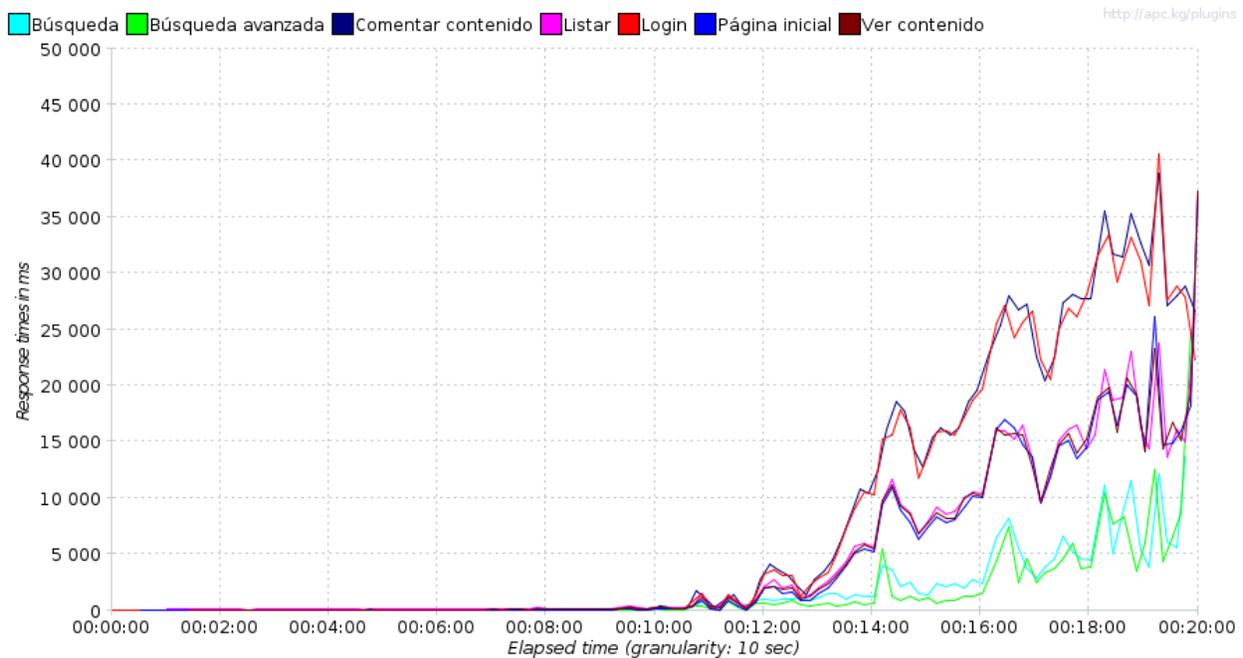


Figura 32 - Tiempos de respuesta: prueba de estrés

Recordando los datos de la **Tabla 8**, el tiempo promedio máximo que los proyectos analizados demoran en responder a una solicitud es de 6.38 segundos, este tiempo es alcanzado por la Didactiteca cuando 791 usuarios realizan simultáneamente peticiones al sistema. Después, el umbral de 10 segundos, descrito por Miller, es alcanzado muy pronto, con 798 usuarios.

Conforme se sigue incrementando el número de usuarios también aumenta el tiempo de respuesta, en especial para las transacciones de inicio de sesión y envío de comentarios. Finalmente el servidor deja de aceptar conexiones cuando el usuario número 1,174 intenta iniciar sesión en el sistema. En este punto el servidor rechaza en promedio 0.3% de las peticiones, lo cual indica que el servidor sigue siendo capaz de procesar una gran cantidad de transacciones, pero el tiempo de respuesta ha dejado de ser tolerable; un usuario común se vería tentado a abandonar el sitio o reenviar su petición, lo cual solo generaría más carga para el servidor. La **Tabla 12** resume este análisis.

Usuarios simultáneos	Tiempo máximo de respuesta [ms]	Observaciones
1 – 300	688	El comportamiento bajo estas condiciones era conocido de antemano.
301 – 640	761	El tiempo de respuesta sigue estable, la diferencia es imperceptible para un usuario.
641 – 790	4,966	Al sistema se le dificulta estabilizarse. Cada nueva conexión supone un incremento en el tiempo de respuesta.
791 – 800	9,884	El tiempo de respuesta sigue en aumento, pero aun dentro de un rango aceptable.
801 – 824	14,404	La Didactiteca comienza a demorar más que cualquiera de los acervos analizados.
825 – 950	30,655	Medio minuto es suficiente para que un usuario comience a frustrarse al utilizar el sistema.
951 – 1,173	77,672	Algunas transacciones demoran más de un minuto, pero el sistema sigue respondiendo a todas.
1,174	-	Empiezan a surgir errores de conexión para el 0.3% de las peticiones.

Tabla 12 - Prueba de estrés: detalles

Gracias a esta prueba se sabe que la Didactiteca es capaz de manejar perfectamente 640 usuarios simultáneos, lo cual supone un aproximado de 2,133 usuarios registrados en el sistema. Además, si el sistema siguiera en expansión, se cuenta con un margen bastante amplio, hasta 800 usuarios conectados simultáneamente, lo cual permitirá contar con el tiempo necesario para optimizar la aplicación conforme a la nueva carga de trabajo.

La **Figura 33** engloba todas las peticiones para denotar el tiempo promedio de respuesta. Considerando tan solo este tiempo es posible ofrecer un margen mayor en donde se considere un buen rendimiento de la Didactiteca, pero esto es debido a que las búsquedas demoran mucho menos que el resto de las transacciones y, por lo tanto, mejoran el tiempo promedio.

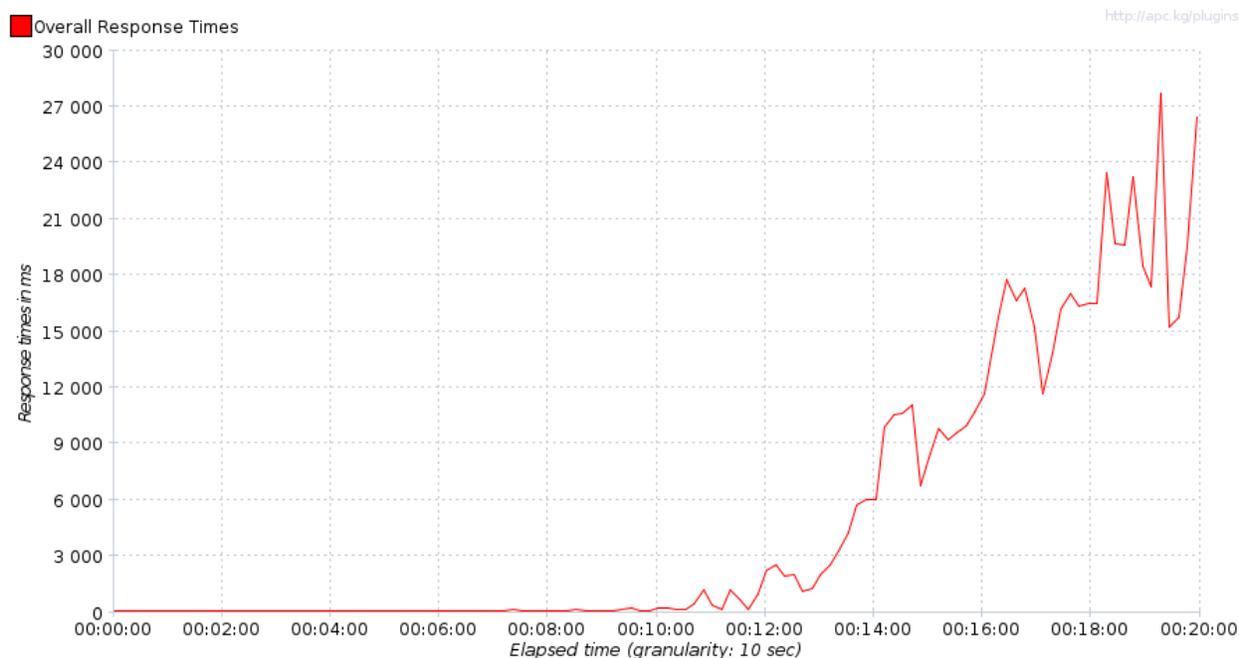


Figura 33 - Tiempo de respuesta promedio: prueba de estrés

La razón de este comportamiento se debe al modo en que se decidió implementar las búsquedas: utilizando índices externos a la base de datos. El índice almacena todo lo relacionado con los recursos educativos de la Didactiteca sobre lo cual se desee realizar una búsqueda.

La implementación de este índice se logró trabajando con el motor de búsqueda Lucene, el cual ofrece ciertas ventajas sobre una búsqueda realizada con lenguaje SQL: (81) (82)

- Puede ordenar los resultados por relevancia, es decir, coloca al inicio los resultados que mejor concuerden con la búsqueda.
- Cuenta con analizadores de texto capaces trabajar con diversos idiomas, entre ellos el portugués y español, para normalizar las búsquedas. Estos analizadores se encargan de filtrar prefijos y sufijos gramaticales, así como algunos tiempos verbales de cada palabra que se desee buscar, lo cual dota de mayor soltura a la búsqueda y permite encontrar resultados más útiles para los usuarios.
- Si el texto ingresado no genera ningún resultado, Lucene puede sugerir al usuario una búsqueda o intentar corregir algún posible error ortográfico basándose en palabras existentes en el índice.
- Un índice externo ofrece protección inherente contra ataques por inyección de código SQL^[19]. Una búsqueda con Lucene no requiere acceder a la base de datos, por lo tanto,

¹⁹ La inyección de código SQL es un ataque en el cual se inserta código malicioso en las cadenas que posteriormente se pasarán a un DBMS para su ejecución. La forma principal de inyección de código SQL consiste en la inserción directa de código en variables especificadas por el usuario que se concatenan con comandos SQL.

cualquier ataque por inyección de código SQL deja de ser un peligro para la aplicación.

Aunado a estas características existe una mejora en los tiempos de respuesta. Esta mejora está garantizada cuando las búsquedas se realizan sobre campos de texto libre cuyo contenido no se relaciona directamente con la estructura de la base de datos, por ejemplo, cuando se requiera buscar un artículo por su título o descripción. Generalmente, cuando un usuario puede buscar información mediante el ingreso de texto se está realizando una búsqueda sobre campos de texto libre. (83)

Para comprobar la ventaja de una búsqueda en Lucene sobre una búsqueda con un DBMS se realizaron pruebas bajo dos contextos:

- Base de datos con índice interno, en donde cada campo sobre el cual se requiera buscar se encuentra indizado con las herramientas que el DBMS, PostgreSQL, dispone.
- Base de datos con índice externo implementado en Lucene, en donde cada campo sobre el cual se requiera buscar ha sido analizado por Lucene para construir el índice.

En ambos contextos se comprobó la respuesta ante el caso de uso *búsqueda rápida*, el cual consiste en que un usuario ingrese el texto que desea buscar y la aplicación se encargará de buscar las palabras en un conjunto de campos referentes a un recurso educativo:

1. Título
2. Nombre del autor
3. Tema
4. Competencia
5. Descripción
6. Palabras clave

La **Tabla 13** muestra el resultado de las pruebas de rendimiento para ambos contextos, en donde es posible observar la ventaja tanto en tiempo de respuesta como en tasa de trabajo que ofrece una búsqueda con un índice externo implementado en Lucene. Ambas pruebas se ejecutaron contemplando 20 usuarios simultáneos realizando una búsqueda cada segundo.

Contexto	Transacciones	Promedio [ms]	Mínimo [ms]	Máximo [ms]	Tasa de trabajo [trans/min]
Búsqueda con DBMS	5,802	182	64	636	996
Búsqueda con Lucene	6,209	105	30	451	1,062

Tabla 13 - Pruebas de rendimiento: búsquedas

6.3 Resultados

El resultado principal de esta tesis fue dotar al Departamento de Portugués del Centro de Enseñanza de Lenguas Extranjeras con un sistema para el apoyo de los cursos presenciales de portugués.

El sistema, referido a lo largo de esta tesis como Didactiteca de Portugués, complementa la nueva propuesta curricular del Departamento del Portugués para la enseñanza del idioma en el CELE. La propuesta implica una reestructuración del plan de estudios y un nuevo enfoque para el aprendizaje del portugués como lengua extranjera, con respecto al antiguo plan de estudios.

Con ayuda de la Didactiteca de Portugués, los profesores pueden crear recursos educativos cuya estructura y objetivos concuerden con la propuesta curricular, puesto que facilita la revisión de calidad al automatizar los procesos de dictaminación.

Cada profesor podrá consultar los materiales desde su hogar o haciendo uso de alguna sala de cómputo en el CELE. El buscar los recursos que necesita deja de ser una tarea meticulosa y propensa a consumir tiempo, bastará con escribir unas cuantas palabras para una búsqueda rápida.

Si la persona, en cambio, prefiere guiarse por la estructura curricular por Tareas o Competencias, bastará con elegir la categoría correspondiente para obtener un listado de recursos educativos. En este listado, si desea visualizar algún recurso podrá hacerlo con tranquilidad y regresar al listado en cualquier momento. Tal comportamiento también está presente en las búsquedas, por lo que un profesor no necesita buscar un mismo término dos veces, simplemente, después de revisar un recurso, podrá regresar a los resultados de la búsqueda.

El sistema es un acervo digital, el cual permite al Departamento de Portugués almacenar, compartir y gestionar un conjunto de recursos educativos, los cuales beneficiarán a cualquier profesor que decida formar parte de la Didactiteca de Portugués, así como a sus alumnos. Los recursos educativos presentes en el sistema se mantendrán en constante actualización ya que los administradores tienen la facilidad de poder solicitar la remoción de cualquier actividad cuya temática haya caducado.

La Didactiteca de Portugués permite que los mismos usuarios califiquen y comenten los recursos educativos que utilizan, aunque el sistema no está pensado para ser una plataforma de diálogo entre los profesores, sí facilita la interacción entre los mismos, lo cual permite tanto críticas constructivas como una oportunidad para que el autor de una actividad pueda retroalimentarse con los comentarios.

A nivel técnico, el sistema proporciona los siguientes servicios:

- **Control de acceso:** el sistema contempla 4 roles, tres de ellos, Miembro, Dictaminador y Administrador, requieren estar registrados en el sistema para poder acceder a los recursos y desempeñar sus funciones. Para ello la Didactiteca contempla formatos de

registro, manejo de sesiones, confirmación de correo electrónico, tiempos de vigencia y restricciones en cuanto a los roles que una persona puede adoptar. El cuarto rol es conocido como Público y engloba todas las acciones que una persona puede realizar sin necesidad de registrarse en el sistema.

- **Resolución automática de dictaminaciones:** dependiendo del resultado que tengan las evaluaciones de ingreso de nuevo material al acervo, el sistema determina automáticamente la siguiente acción a realizar:
 - Enviar de la actividad a los administradores para su publicación.
 - Devolver la actividad al autor para una corrección.
 - Pedir la evaluación de un nuevo dictaminador en caso de haber discrepancias significativas y dar seguimiento a dicha evaluación.
- **Notificaciones por correo electrónico:** varias de las acciones que pueden realizarse en la Didactiteca de Portugués cuentan con un tiempo máximo (días, semanas, etc.) en el cual se espera una respuesta, el sistema notifica a los usuarios con anticipación acerca de la acción pendiente.
- **Almacenamiento y eliminación automática de eventos temporales:** En la página principal del sitio web, un usuario puede consultar las conferencias, exposiciones y demás eventos que los administradores publiquen. Una vez pasada la fecha del evento este queda automáticamente fuera de la vista de los usuarios, pero un administrador todavía puede consultar el listado de todos los eventos que se han publicado a lo largo de un año, lo cual facilita el mantener un registro de los mismos. Después de que ha pasado el tiempo suficiente, el sistema eliminará paulatinamente los eventos más antiguos. Todo ello sin intervención alguna de los administradores.
- **Almacenamiento y visualización gráfica de información estadística:** la Didactiteca de Portugués guarda información acerca de las visitas, comentarios y descargas que cada recurso educativo va generando. Toda la información está disponible a los administradores de modo visual, a través de gráficas en donde es posible visualizar los datos de en un rango de tiempo flexible.
- **Creación de una versión imprimible para los metadatos de un recurso educativo:** Si un profesor desea imprimir una ficha con los datos que describen una actividad, puede solicitar al sistema la creación de una versión en formato PDF. La creación de este documento se realiza de manera dinámica, por lo que no requiere de espacio extra en disco duro.

CONCLUSIONES

La realización de esta tesis ha permitido que el Departamento de Portugués cuente ahora con un sistema en donde los profesores pueden publicar y consultar recursos educativos para los cursos presenciales de portugués en el CELE.

Debido a que durante la realización de la tesis se ha utilizado el estándar de ingeniería de software ISO/IEC 29110, con especial énfasis en el proceso de implementación de software, se aseguró que la definición de los requerimientos consolidara una descripción completa del sistema que el Departamento de Portugués necesita. Gracias a ello, la Didactiteca de Portugués se encuentra estrechamente ligada con la actualización curricular de departamento.

Las pruebas de rendimiento realizadas sobre el sistema permitieron asegurar su buen funcionamiento, tanto bajo condiciones normales de operación como bajo una enorme carga de trabajo. También permitieron contemplar una posible expansión de la Didactiteca de Portugués al simular una carga de trabajo equivalente a la que se tendría con un mayor número de usuarios registrados en el sistema.

En lo referente a los procedimientos y medidas a considerar durante la realización se las pruebas de rendimiento, se realizó una investigación cuyos resultados fueron completamente inesperados para mí, puesto que, como se explica en el capítulo 6.2.2, el campo de las pruebas de software actualmente no cuenta con algún estándar ampliamente reconocido. Diversas instituciones y empresas^[20] ofrecen certificaciones, con sus respectivos cursos, para personas que deseen profundizar dentro del tema de las pruebas de software, lo cual sugiere que pueda existir al menos una línea base sobre la cual una persona pueda guiarse al momento de desarrollar las pruebas de software. Sin embargo, dado que una certificación tiene un costo, es imposible conocer a detalle el contenido de un curso de certificación. Para aquellos cursos de los cuales se puede tener noción de su contenido (84) (85) (86), se observa que existen dos casos:

- No se hace mención alguna de estándares en lo referente a pruebas.
- Se mencionan las pruebas de *benchmarking* categorizadas como estándares.

Una prueba de *benchmarking* es una prueba comparativa para medir el rendimiento de un sistema o componente (73). La comparación se realiza al aplicar un mismo proceso de prueba a un conjunto de sistemas o componentes de un equipo (CPU, GPU, disco duro, etc.), con ayuda de una herramienta de software. Entonces, si se desea probar, por ejemplo, el rendimiento de un CPU, se utiliza una herramienta de software que ejecuta el proceso de pruebas en el equipo que contenga el CPU a ser analizado. Esta herramienta, idealmente, también contará con un listado de distintos CPUs con los cuales se realizará la comparativa, puesto que ya han sido analizados anteriormente por los mismos desarrolladores de la herramienta de software.

²⁰ El sitio de Wikipedia ofrece un listado de programas de certificación para pruebas de software: http://en.wikipedia.org/wiki/Software_testing#Certifications

Como puede observarse una prueba de *benchmarking* ofrece una comparativa contra otros elementos similares. Si el componente de software o hardware que se desea probar puede ser comparado contra un componente ampliamente reconocido o utilizado en el mercado, entonces es posible asegurar un margen más realista de su funcionamiento. Pero cabe destacar que cada herramienta de *benchmarking* contiene procesos de prueba distintos y, por lo tanto, no estándares.

Conforme a la decisión de realizar una implementación propia en vez de modificar y configurar un CMS existente, es posible resaltar los siguientes aspectos en favor de la implementación propia:

- Libertad para elegir tanto el lenguaje de programación como la plataforma de desarrollo: para el caso de esta tesis se escogió una plataforma cuyos paradigmas está estrechamente relacionados tanto con la programación orientada a objetos como con la arquitectura MVC, lo cual permitió un desarrollo más ágil de los casos de uso contemplados.
- Las limitantes en cuanto a la implementación son menores: si se requiere implementar algún elemento demasiado especializado requerirá de un mayor esfuerzo para el caso de un CMS, principalmente con cuestiones de compatibilidad.
- Mayor potencial para la investigación académica: dado que un CMS proporciona tanto funciones básicas como algunas más avanzadas, no se requerirá profundizar en varios aspectos durante el desarrollo. Por otro lado, un desarrollo propio es idóneo si se desea aprender cómo es que se realiza un proyecto de la magnitud del sistema resultante de esta tesis.
- Mayor configuración y personalización: todo aspecto configurable dependerá tan solo de la plataforma de desarrollo y del sistema mismo que se programe, por lo tanto, es posible modificar hasta el aspecto más mínimo del sistema. Con ello es posible asegurar un sistema hecho a la medida.

Por otro lado, un desarrollo propio incluye también sus propias desventajas:

- El sistema debe ser creado desde sus cimientos: Esta es la mayor desventaja. Un CMS provee una base sólida sobre la cual trabajar, por medio de configuraciones y añadidura de código propio, lo cual libera al programador de una carga considerable.
- El cumplimiento de estándares es responsabilidad total del desarrollador: Algunos CMS cuentan con elementos estandarizados, por ejemplo, DSpace implementa el estándar de metadatos Dublin Core. Esto ayuda a los desarrolladores a implementar un sistema con mayores facilidades de interoperabilidad.

Para el caso específico de esta tesis, las ventajas de un desarrollo propio resultaron ser un factor de gran importancia, por lo cual las desventajas pudieron ser sobrellevadas sin grandes complicaciones.

Personalmente, la realización de esta tesis fue una gran oportunidad para indagar acerca de la realización de un proyecto formal de software, desde la toma requerimientos, la definición y redacción de los documentos implicados hasta la implementación misma del sistema.

Durante la fase de análisis y diseño se requirieron conocimientos de:

- Ingeniería de software
- Modelado de sistemas, para lo cual se utilizó UML.
- Modelado de datos, en donde las estructuras y sus relaciones fueron modeladas en diagramas Entidad-Relación.

La implementación requirió de conocimientos en una variedad de lenguajes de programación, así como otros aspectos técnicos:

- Lenguajes de programación:
 - Java
 - Groovy
 - JavaScript
- Plataforma de desarrollo Grails.
- Base de datos PostgreSQL.
- Presentación de información y diseño gráfico:
 - HTML
 - CSS

Una de las características más interesantes de la Didactiteca de Portugués es que está completamente orientada a los usuarios. Son sus aportes los que contribuyen a crecimiento y prosperidad del sistema, con lo cual la Didactiteca de Portugués cae dentro de la categoría de Web 2.0.

El sistema incita a los usuarios a contribuir para poder beneficiarse con las aportaciones de todos los demás, y, en cierto modo, también los obliga a ello, puesto que tales contribuciones aseguran que la cuenta del usuario en el sistema permanezca vigente hasta un máximo de 6 meses sin que se requiera de una nueva contribución.

Conforme la Didactiteca de Portugués vaya creciendo los profesores contarán con cada vez más recursos educativos para sus clases presenciales.

Actualmente, la computación se ha convertido en un elemento distintivo del progreso. Para el caso del Departamento de Portugués, su uso significará tener disponibles en línea recursos didácticos para sus clases presenciales, junto con todas las demás ventajas que se han descrito a lo largo de la tesis.

Conforme la Didactiteca de Portugués alcance una mayor difusión aportará mayores beneficios para sus usuarios, no solo porque existirán más recursos educativos, sino también porque el Departamento de Portugués del CELE tendrá una mejor oportunidad para seguir actualizándose, ya que todo usuario registrado tiene la libertad de aportar comentarios y calificar los recursos que ha utilizado.

Este tipo de comportamiento colaborativo, en donde los usuarios y el sistema se ven beneficiados mutuamente, suele serle familiar a quienes se dedican al estudio e implementación de sistemas computacionales.

Por ejemplo, el sitio web de *Stack Overflow*^[21] permite a sus usuarios realizar preguntas de carácter técnico, principalmente de programación, en donde son los mismos usuarios quienes también responden las preguntas; se cuenta con la posibilidad de comentar y calificar tanto las preguntas como las respuestas. Además, son los mismos usuarios quienes moderan y gestionan gran parte del sitio.

Como puede observarse, la Didactiteca de Portugués guarda cierta similitud con algunos aspectos de Stack Overflow, e incluso de otros sistemas colaborativos. Puesto que es una creciente tendencia, contar con entornos que permitan a los usuarios compartir su conocimiento al mismo tiempo que se ven beneficiados por formar parte una comunidad.

Todo ello supone un avance en cuanto a la enseñanza de un idioma, que ha empezado en el Departamento de Portugués del CELE, pero podría ser adoptado por alguna otra institución.

²¹ <http://stackoverflow.com/>

TRABAJO A FUTURO

8.1 Configuración flexible

El sistema está creado para acoplarse a las necesidades específicas del Departamento de Portugués, pero no por ello se ha dejado a un lado la flexibilidad. Por ejemplo, la taxonomía básica de los recursos educativos está englobada en dos categorías: Competencias y Tareas (para mayor detalle, favor de referirse al capítulo 4.1), las cuales ya han sido definidas por el Departamento de Portugués, sin embargo el sistema los utiliza de forma dinámica. Pueden existir más Tareas o menos Competencias y tal hecho no impactaría de manera alguna el comportamiento del sistema, solo podría requerir alguna actualización en cuanto al diseño visual.

Esta flexibilidad está presente en gran parte de los componentes de la Didactiteca de Portugués, por lo tanto es posible realizar ajustes y adaptaciones sin necesidad de modificar la lógica interna del sistema. El elemento faltante sería un punto de interacción para que un administrador pueda realizar tales modificaciones, el cual no está contemplado dentro de los requerimientos, pero incrementaría la autosuficiencia del sistema, ya que los mismos administradores podrían configurar los aspectos básicos del sistema.

Este tipo de configuración se tiene en mente debido al estado actual de la propuesta curricular del Departamento de Portugués. Dado que el sistema entrará plenamente en funcionamiento junto con la actualización curricular, cualquier cambio que se realice al plan de estudios debe reflejarse en el sistema, ya que los recursos didácticos que la Didactiteca aloja están estrechamente ligados con el plan de estudios.

8.2 Visualización de materiales en el sitio

Cada recurso educativo debe contar con al menos un material, el cual corresponde a la descripción de la actividad que los alumnos deben realizar en clase. Este material será un documento escrito, pero no es lo único que puede contener un recurso educativo, también es posible que existan imágenes, videos, audio, etc. La posibilidad de visualizar cada material sin necesidad de descargarlo volvería más amena la experiencia de los usuarios de la Didactiteca de Portugués.

Esta mejora requerirá la implementación de nuevos elementos dentro del sistema que permitan visualizar los formatos más populares, especialmente archivos de audio y video. Para ello se contemplan dos vías de diseño e implementación:

1. Integrar la visualización directamente dentro del sistema.
2. Delegar la visualización a sitios web externos.

Para el primer caso, se requiere analizar qué formatos de archivo son los más utilizados para implementar los reproductores necesarios en el sistema. Bajo este contexto será responsabilidad de los usuarios el utilizar los formatos que la Didactiteca de Portugués acepte. La posibilidad de descargar los materiales siempre estará presente, tanto para formatos reproducibles en el sitio como para cualquier otro.

Dado que es imposible conocer el uso final que se le dará a los materiales, se debe tener especial cuidado para cualquier implementación que implique la conversión de los archivos descargables a otro formato. Convertir los archivos garantizaría que una mayor cantidad de materiales fueran reproducibles en el sitio, pero podría dificultar su uso, una vez que sea descargado.

Para el segundo caso, se requiere analizar cuáles son los sitios que permiten un manejo más fácil de las visualizaciones, desde el punto de vista de la implementación.

Dentro de los proyectos analizados, el MIT OCW resulta un perfecto ejemplo, además, es el único que implementa la visualización de material audiovisual en el sitio, aunque el proceso parece implicar un mayor esfuerzo por parte del autor o administradores.

Cuando un curso en el MIT OCW contiene material audiovisual se proporciona un *reproductor Flash*^[22] en el sitio, tal como se muestra en la **Figura 34**. Pero dicho material no se encuentra en el sistema del MIT OCW, a diferencia de los documentos escritos o imágenes, sino que está enlazado desde una página en el sitio de *YouTube*^[23] para la visualización en el reproductor Flash. Mientras que los enlaces para la descarga del material se encuentran disponibles desde dos ubicaciones:

²² Un reproductor Flash es un elemento de software que permite la reproducción de contenido audiovisual en el formato propietario *Flash Video*. Para ejecutar un reproductor Flash en una página web, el usuario requiere que su explorador cuente con una versión compatible de *Adobe Flash Player*. En el sitio oficial de Adobe se puede obtener mayor información al respecto: <http://www.adobe.com/es/products/flashplayer.html>

²³ La página actual donde se pueden encontrar los videos del MIT OCW: <http://www.youtube.com/user/MIT/>

- El portal de *iTunes U*^[24], donde los videos son accesibles únicamente desde el reproductor multimedia *iTunes*, ya sea en una computadora o cualquier dispositivo portátil que cuente con iTunes.
- La biblioteca digital *Internet Archive*^[25], donde una persona puede descargar el video en formato MP4 y reproducirlo más tarde.

Home > Courses > Electrical Engineering and Computer Science > Performance Engineering of Software Systems > Video Lectures > Lecture 1: Matrix Multiply: A Case Study

Lecture 1: Matrix Multiply: A Case Study



00:00 00:00

next >>

About this Video Playlist Lecture Notes Download this Video

Free Downloads

Video

[iTunes U \(MP4 - 153MB\)](#)

[Internet Archive \(MP4 - 153MB\)](#)

Figura 34 - Reproducción de video en el MIT OCW

²⁴ Una descripción de iTunes U para profesores puede leerse en el sitio oficial de Apple:
<http://www.apple.com/education/itunes-u/>

²⁵ <http://archive.org/>

Implementar la visualización de material audiovisual del modo en que el MIT OCW lo realiza implica crear copias del archivo original en sitios externos, restringiendo enormemente los formatos a utilizar, pero reduciendo tanto la carga de trabajo para el sistema y minimizando en trabajo que se requiere para el diseño y desarrollo.

Este apartado requerirá de un análisis más extenso, para contemplar las vías de implementación existentes, así como sus ventajas y desventajas.

8.3 Elementos de difusión

Recordando el análisis expuesto en el capítulo 4.2, el modo de difusión más efectivo suele ser el que llevan a cabo los usuarios que ya han comprobado las ventajas de utilizar cierto sistema.

Aproximadamente 25.4% de los usuarios de un acervo se informaron del mismo a través de un colega (41). Lo cual nos indica que, independientemente de los esfuerzos realizados por las entidades educativas o cualquiera que sea el organismo encargado, los usuarios tienen una gran influencia en cuanto a la difusión que tenga un acervo.

La Didactiteca de Portugués actualmente no contempla el uso de elementos de difusión que permitan a los usuarios dar a conocer los recursos con que cuenta el sistema, incluyendo aquellos que son de su autoría.

Integrar elementos que ayuden a los usuarios a promover la Didactiteca de Portugués, permitiría una difusión más efectiva y sencilla de la que pudiera efectuar cualquier usuario por su parte. Por ejemplo, si una persona quisiera enviar por correo electrónico un enlace directo hacia cualquier recurso, se vería en la necesidad de copiarlo de su propio navegador, tal acción es propensa a errores.

Retomando los proyectos analizados en el capítulo 4.3, se observó que tres de ellos cuentan con algún elemento que facilite la difusión de un recurso educativo. Sorpresivamente, cada uno de estos proyectos utiliza solo un medio de difusión y todos son distintos entre sí:

- MIT OCW: ofrece la posibilidad de enviar un correo electrónico a una o varias personas. El envío se hace por medio de la cuenta de correo del MIT OCW, colocando, tanto en el título del correo como en el remitente, la dirección de correo electrónico de quien compartió el recurso (37). La **Figura 35** resalta la opción que el sitio ofrece para compartir el recurso educativo haciendo uso del correo electrónico, mientras que la **Figura 36** muestra la ventana emergente que se proporciona para esta opción.

- › Course Home
- › Syllabus
- › Calendar
- › Course Notes
- › Video Lectures
- › Assignments
- › Exams

- › Download Course Materials

- › Send us your feedback
- › Cite this course
- › Email this page
- › Newsletter sign-up
- › Donate

Discrete Stochastic Processes

As taught in: Spring 2011

Tandem queues: A stable M/M/1 queue has a Poisson output at the input rate. The next queue also has a Poisson output at that rate. (Image by MIT OpenCourseWare, adapted from Prof. Robert Gallager's course notes.)

Instructors:
Prof. Robert Gallager

MIT Course Number:
6.262

Level:
Graduate

Course Features

- › Video lectures
- › Exams and solutions
- › Assignments and solutions

Figura 35 - MIT OCW: enviar por email

MITOPENCOURSEWARE
 MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Send friends a link to this web page and a short message by completing this form.

***To** (If more than one, separate with commas)

***Your Email Address**

Copy me on this email

Subject:
MIT OpenCourseWare | Electrical Engineering and Comp

Message: (optional)

I saw this page on the MIT OpenCourseWare site and thought you might be interested in this:

*required fields

Memorizys

aground

Escribe las dos palabras:

stop spam.
read books.

We will not record or use the e-mail address you supply for anything other than sending your message. For more information see our [Privacy Policy](#)

Figura 36 - MIT OCW: forma para envío de email

- CREA: Utiliza la difusión por la red social Twitter^[26] para compartir un enlace hacia el recurso educativo. La **Figura 37** resalta la opción que el sitio ofrece para compartir un recurso educativo por Twitter, mientras que la **Figura 38** muestra el mensaje que se genera al utilizar tal opción.

The screenshot shows a sidebar on the left with navigation options: 'Materiales destacados', 'Creativos trabajando', 'Comité Técnico', 'Preguntas frecuentes', 'Explorar', 'Formación', 'Áreas del conocimiento', 'Títulos', 'Autores', 'Por fecha', 'Búsqueda avanzada', 'Entrar a', and 'Mi CREA usuarios autorizados'. The main content area displays a resource card with the following details:

- Twittear OA** (button)
- Título:** Categorías gramaticales del español
- Autor:** Trejo González, Hugo; Ovarhage Chávez, Gerlind
- Palabras clave:** Gramática generativa, Escritura, Español, Ortografía
- Fecha de publicación:** 19-Jun-2012

Figura 37 - CREA: compartir por Twitter

The screenshot shows the Twitter sharing interface. At the top, there is a Twitter logo and a search bar. Below that, the text 'Comparte un link con tus seguidores' is displayed. A text input field contains the URL 'http://www.crea.udg.mx/handle/123456789/362'. To the right of the input field, there is a '119' and a blue 'Twittear' button.

Figura 38 - CREA: redactar mensaje en Twitter

²⁶ Twitter es un servicio que permite a sus usuarios publicar mensajes breves. Los mensajes son enviados a las personas que se han suscrito a las actualizaciones del usuario en cuestión. Actualmente es una de las redes sociales más populares. La página principal de Twitter se encuentra en: <https://twitter.com/>

- MERLOT: No proporciona alguna herramienta que automatice la difusión en redes sociales o correo electrónico, sin embargo, cada recurso dispone de un código en formato *Quick Response (QR)*^[27] que contiene el enlace al recurso educativo. La **Figura 39** resalta la ubicación del código QR en el sitio de MERLOT. Generalmente, una persona utilizará un dispositivo portátil para escanear el código y poder compartir un enlace al recurso educativo por el medio que crea conveniente, mediante alguna aplicación de software adicional.

Image Enhancement Algorithms

No Image Available

Location: [Go to Material](#)

Material Type: [Simulation](#)

Date Added to MERLOT: febrero 06, 1998

Date Modified in MERLOT: marzo 31, 2003

[\[Report Broken Link For This Material\]](#)

Author: [Jauvane C. de Oliveira](#)

Submitter : [Jason Isaac](#)

Description:
This is an applet that performs a collection of Image Enhancement algorithms (Image Processing)

Browse in Categories:
- [Science and Technology/Engineering](#)

More information about this material:
Primary Audience: College General Ed
Mobile Compatibility: Not specified at this time
Language: English
Cost Involved: no
Source Code Available: yes
Accessibility Information Available: no
Copyright: yes
Creative Commons: unsure

About this material:
Peer Reviews (not reviewed)
Workflow status (Not triaged)
Comments (none)
Learning Exercises (none)
Personal Collections (none)
Accessibility Info (none)

Add your own:
[Write a comment](#)
[Create a learning exercise](#)
[Add accessibility information](#)

Add to a personal collection:

[Report this as an Inappropriate Material](#)

[What's This?](#)

Figura 39 - MERLOT: código QR

Para integrar herramientas de difusión en el sitio de la Didactiteca de Portugués, se requiere preguntar a los usuarios acerca de los medios digitales que comúnmente utilizan para compartir información en la red.

Además se debe contemplar el carácter cerrado del sistema, en donde solo usuarios registrados pueden ver un recurso completo. Una opción viable sería proporcionar enlaces directos a los recursos, puesto que el sistema se encarga de entregar, automáticamente, la vista que corresponde al rol de la persona. Por otro lado podría optarse por compartir únicamente un conjunto de metadatos y un enlace a la Didactiteca.

²⁷ Un código Quick Response, es un código matricial que permite el almacenamiento de información. Su uso se ha popularizado junto con el creciente mercado de dispositivos móviles con cámara, con los cuales se facilita la lectura de estos códigos.

APÉNDICES

9.1 Control de versiones

El control de versiones es un sistema que registra los cambios en un archivo o conjunto de archivos a través del tiempo para que se puedan recuperar versiones específicas más tarde (87).

Un sistema de control de versiones cuenta con varias características: (88)

- El mantenimiento versiones o revisiones antiguas de archivos.
- Capacidad de contar con información asociada a cada una de las versiones o revisiones.
- La posibilidad de que varios usuarios estén trabajando sobre un mismo archivo y se efectúen sin bloqueos las modificaciones realizadas por cada uno de los usuarios,
- Facilita la comparación de diferencias entre archivos de distintas versiones.
- Funge como un almacén de datos que contiene la información de un proyecto a lo largo del tiempo.

Al momento de iniciar el proyecto de tesis, el CAD contaba con el sistema Subversion (SVN)^[28], el cual permite un control de versiones centralizado. Esto significa que, en SVN, se tienen un único servidor que contiene todos los archivos versionados, y un cliente consulta los archivos de ese lugar central, tal como se expone en la **Figura 40**.

Los archivos de la Didactiteca de Portugués, tanto código como documentación, se gestionaban dentro de un servidor SVN ubicado dentro de la red del CAD.

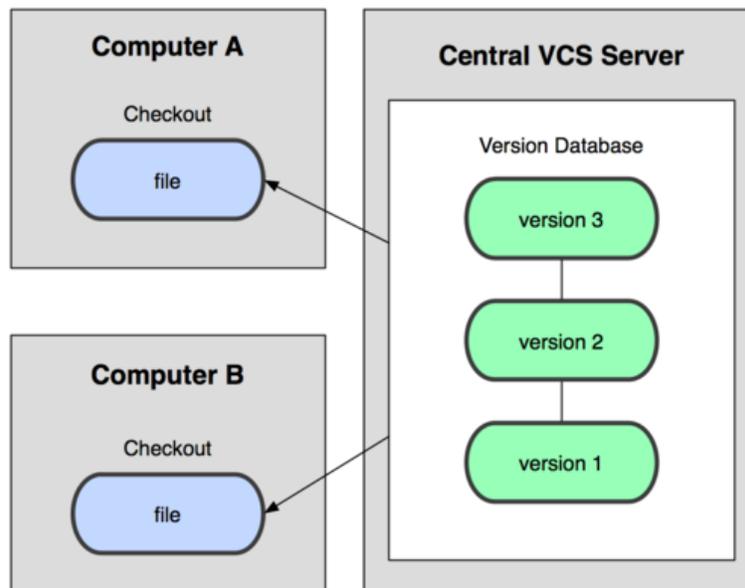


Figura 40 - Control de versiones centralizado (87)

²⁸ <http://subversion.apache.org/>

Más tarde, durante la fase de implementación de la tesis, el proyecto se actualizó a Git^[29], el cual es un sistema de control de versiones distribuido. Con Git cada cliente cuenta con una copia local completa del proyecto, por lo que no se requiere una conexión al servidor para guardar cambios o revisar versiones anteriores. El comportamiento básico de Git se puede observar en la **Figura 41**.

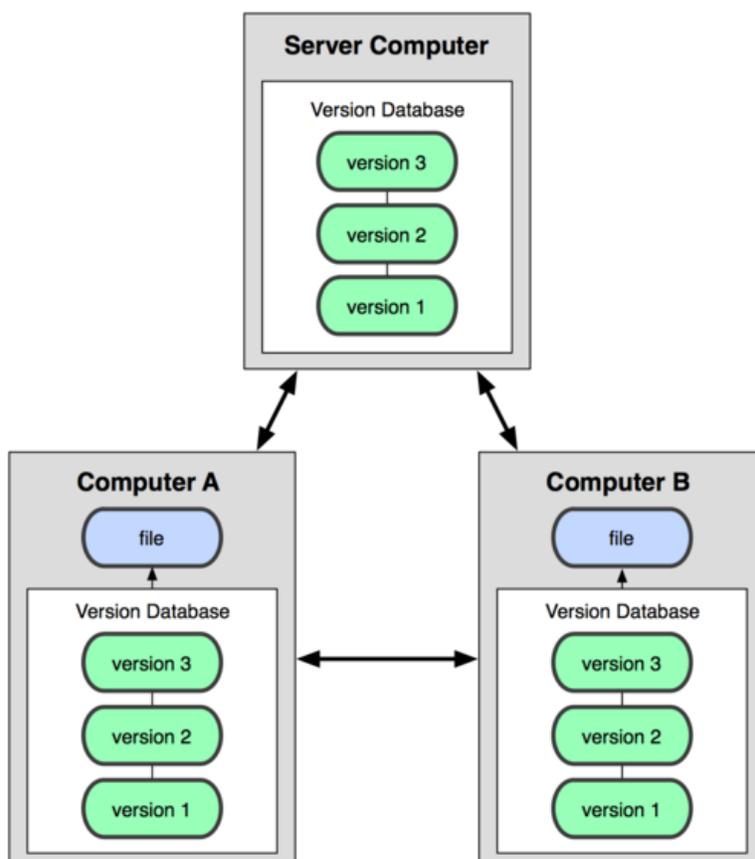


Figura 41 - Control de versiones distribuido (87)

Utilizar un control de versiones también facilita solucionar problemas dentro del código de una aplicación, si es que tal problema surge al modificar algún archivo. Es posible revisar de manera automática diversas versiones para encontrar aquella en donde el problema se generó. Esto es particularmente importante cuando el equipo de trabajo es grande y los cambios en el código se realizan de manera constante. (87)

El uso de un sistema para control de versiones, ya sea centralizado o distribuido, permite una mejor gestión, tanto del código de una aplicación como de la documentación asociada. Además esto asegura el cumplimiento de dos aspectos relevantes de la norma ISO/IEC 29110:

1. Contar con un repositorio para el proyecto.
2. Utilizar una estrategia de control de versiones.

²⁹ <http://git-scm.com/>

BIBLIOGRAFÍA

1. **Titone, Renzo.** *Metodología didáctica*. Madrid : Ediciones Rialp S.A., 1981.
2. **Fenstermacher, Gary y Soltis, Jonas.** *Enfoques de la enseñanza*. Buenos Aires : Amorrortu, 1999.
3. **Stenhouse, Lawrence.** *Investigación y desarrollo del curriculum*. Londres : Ediciones Morata, 1991.
4. **Dávila, Sergio.** El aprendizaje significativo: Esa extraña expresión (utilizada por todos y comprendida por pocos). [DePa] Departamento de Programas Audiovisuales. [En línea] [Citado el: 18 de marzo de 2012.]
http://depa.fquim.unam.mx/amyd/archivero/AUSUBELAPRENDIZAJESIGNIFICATIVO_1677.pdf.
5. **Mills Gagné, Robert.** *Las condiciones del aprendizaje*. Ciudad de México : Interamericana, 1987.
6. **Ávila, Patricia.** Aprendizaje con nuevas tecnologías: paradigma emergente. *Investigación y Desarrollo - ILCE*. [En línea] [Citado el: 18 de marzo de 2012.]
<http://www.cite.com.mx/documentos/antologia/iiia.pdf>.
7. **Fundación Centro de Investigación de Economía y Sociedad CIES .** Guía metodológica para la creación y adaptación de cursos para la formación e-learning. [En línea] [Citado el: 14 de agosto de 2012.] http://www.grupcies.com/fileadmin/CIESFilesAll1/pdf-Privados/6Guia_elearning_castellano_01.pdf.
8. **Barceló, María.** Hacia una economía del conocimiento. [En línea] 2001. [Citado el: 17 de agosto de 2012.] <http://books.google.com.mx/books?id=jn3FdhLNuBIC>.
9. **Tiffin, John y Rajasingham, Lalita.** *En busca de la clase virtual: la educación en la sociedad de la información*. Barcelona : Editorial Paidós, 1997.
10. **Casamayor, Gregorio.** *La formación on-line*. Barcelona : Grao, 2008.
11. **Shatz, David.** *Peer review: a critical inquiry*. Lanham : Rowman & Littlefield, 2004.
12. **Giordanino, Eduardo.** Sistemas de evaluación de trabajos para publicaciones científicas (peer review). *Ciencia y Técnica Administrativa*. [En línea] 2005. [Citado el: 24 de marzo de 2012.] <http://www.cyta.com.ar/ta0402/v4n2a1.htm>.
13. **Rodríguez, Arturo.** La gestión del conocimiento científico-técnico en la universidad: un caso y un proyecto. *ADDI: Repositorio Institucional de la Universidad del País Vasco*. [En línea] 2001. [Citado el: 24 de marzo de 2012.]
https://addi.ehu.es/bitstream/10810/6883/1/CdG_111.pdf.
14. **Steib, Nicolas.** *Gestión del Conocimiento: algo más que información*. s.l. : Training & Development Digest, 1999.
15. **Anumba, Chimay y Egbu, Charles.** *Knowledge Management in Construction*. Chichester : John Wiley and Sons, 2005.
16. **García-Tapial, Joaquín.** Gestión del conocimiento y empresa. Una aproximación a la realidad española. [En línea] Escuela de Organización Industrial EOI, 2002. [Citado el: 19 de agosto de 2012.] http://media.eoi.es/nw/Multimedia/PublicacionesEOI/2002_Libro_52.pdf.
17. **Akgün, Ali y Lynn, Gary.** *Organizational learning: a socio-cognitive framework*. s.l. : Human Relations, 2003.
18. **Centro de Investigación sobre la Sociedad del Conocimiento.** *Modelo Intellectus: medición y gestión del capital intelectual*. Madrid : Universidad Autónoma de Madrid, 2003.

19. **Ordóñez, Patricia.** Importancia estratégica de la medición del capital intelectual en las organizaciones. [En línea] [Citado el: 6 de agosto de 2012.] http://www.injef.com/revista/empresas/pop_991217.htm.
20. **Brooking, Annie.** *Intellectual Capital: Core Asset for the Third Millennium Enterprise.* Londres : Inter-national Thomson Business Press, 1996.
21. **Crow, Raym.** A Guide to Institutional Repository Software. *Open Society Foundations.* [En línea] 2004. [Citado el: 3 de agosto de 2012.] Disponible en: http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf.
22. **Lynch, Clifford.** Digital Collections, Digital Libraries, and the Digitization of Cultural Heritage Information. *First Monday.* [En línea] 2002. [Citado el: 1 de agosto de 2012.] <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/949/870>.
23. **McDowell, Cat.** Evaluating Institutional Repository Deployment in American Academe Since Early 2005. *D-Lib Magazine.* septiembre/octubre, 2007, Vol. 13.
24. **Crow, Raym.** *The Case for Institutional Repositories: A SPARC Position Paper.* Washington : Scholarly Publishing & Academic Resources Coalition, 2002.
25. **Secker, Jane.** *Electronic resources in the virtual learning environment.* Oxford : Chandos, 2004.
26. **Camps, Rafael.** *Bases de datos.* Cataluña : Editorial UOC, 2002.
27. **Cobo, Angel.** *Diseño y programación de bases de datos.* Madrid : Editorial Visión Libros, 2007.
28. **Rob, Peter y Coronel, Carlos.** *Sistemas de bases de datos: Diseño, implementación y administración.* Ciudad de México : Cengage Learning Editores, 2003.
29. **Senso , José.** El concepto de metadato. Algo más que descripción de recursos electrónicos. [En línea] 2003. [Citado el: 11 de agosto de 2012.] <http://www.scielo.br/pdf/ci/v32n2/17038.pdf>.
30. **Bruguera, Enric y Lara, Pablo.** *Los blogs y La organización del conocimiento.* Cataluña : Editorial UOC, 2007.
31. **Iannella, Renato Iannella y Waugh, Andrew.** Metadata: Enabling the Internet. [En línea] [Citado el: 11 de agosto de 2012.] <http://archive.ifla.org/documents/libraries/cataloging/metadata/ianr1.pdf>.
32. **Durán, Francisco, Gutiérrez, Francisco y Pime, Ernesto.** *Programación orientada a objetos con Java.* Madrid : Thomson-Paraninfo, 2007.
33. **University of Nottingham.** *The Directory of Open Access Repositories - OpenDOAR.* [En línea] [Citado el: 16 de agosto de 2012.] <http://www.opendoar.org/index.html>.
34. **Carpenter, Hutch.** Tim O'Reilly Course Corrects the Definition of Web 2.0. [En línea] 2008. [Citado el: 10 de agosto de 2012.] <http://bhc3.com/2008/10/01/tim-oreilly-course-corrects-the-definition-of-web-20/>.
35. **Ribes, Xavier.** La Web 2.0. El valor de los metadatos y de la inteligencia colectiva. *TELOS.* [En línea] 2007. [Citado el: 10 de agosto de 2012.] <http://sociedadinformacion.fundacion.telefonica.com/telos/articuloperspectiva.asp@idarticulo%3D2&rev%3D73.htm>.
36. **de la Torre, Aníbal.** Definición de Web 2.0. [En línea] 2006. [Citado el: 10 de agosto de 2012.] http://www.adelat.org/index.php?title=conceptos_clave_en_la_web_2_0_y_iii.
37. **Massachusetts Institute of Technology.** *MIT OpenCourseWare.* [En línea] [Citado el: 15 de agosto de 2012.] <http://ocw.mit.edu/index.htm>.
38. **University of Southampton.** *Registry of Open Access Repositories.* [En línea] [Citado el: 16 de agosto de 2012.] <http://roar.eprints.org/>.

39. **Keefner, Alice.** Los repositorios digitales universitarios y los autores. *DIGITUM*. [En línea] 2007. [Citado el: 16 de agosto de 2012.] <http://digitum.um.es/jspui/bitstream/10201/4018/3/1201.pdf>.
40. **Santos-Hermosa, Gema, Ferran-Ferrer, Núria y Abadal, Ernest.** Recursos educativos abiertos: Repositorios y uso. [En línea] 2012. [Citado el: 16 de agosto de 2012.] <http://accesoabierto.net/sites/accesoabierto.net/files/Santos-Ferran-Abadal-EPI.pdf>.
41. **Swan, Alma y Brown, Sheridan.** Open access self-archiving: An author study. *University of Southampton*. [En línea] 2005. [Citado el: 15 de agosto de 2012.] <http://eprints.soton.ac.uk/id/eprint/260999>.
42. **UNESCO.** Global Open Access Portal. [En línea] [Citado el: 13 de agosto de 2012.] <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/>.
43. **Instituto Nacional de Salud Pública.** *Artemisa en Línea*. [En línea] [Citado el: 17 de agosto de 2012.] <http://www.artemisaenlinea.org.mx/>.
44. **Universidad del Claustro de Sor Juana.** Repositorio del Claustro de Sor Juana. *Biblioteca Sor Juana Inés de la Cruz*. [En línea] [Citado el: 17 de agosto de 2012.] <http://201.147.150.252:8080/jspui/>.
45. **Universidad de Guadalajara.** *Centro de Recursos para la Enseñanza y el Aprendizaje*. [En línea] [Citado el: 15 de agosto de 2012.] <http://www.crea.udg.mx/index.jsp>.
46. **Tecnológico de Monterrey.** *DAR (Desarrolla, Aprende y Reutiliza)*. [En línea] [Citado el: 16 de agosto de 2012.] <http://catedra.ruv.itesm.mx/>.
47. **Tecnológico de Estudios Superiores de Occidente.** *EduDoc Centro de Documentación sobre Educación*. [En línea] [Citado el: 19 de agosto de 2012.] <http://www.fronteraseducativas.com/edudoc/>.
48. **Universidad Nacional Autónoma de México.** Repositorio de la Facultad de Filosofía y Letras. UNAM. [En línea] [Citado el: 10 de agosto de 2012.] <http://ru.ffyl.unam.mx:8080/jspui/>.
49. **OpenCourseWare Consortium Staff.** OpenCourseWare Consortium. [En línea] [Citado el: 19 de agosto de 2012.] <http://www.ocwconsortium.org/>.
50. **California State University.** *MERLOT*. [En línea] [Citado el: 17 de agosto de 2012.] <http://www.merlot.org/merlot/index.htm>.
51. **Universidad de los Andes.** *Funes: repositorio digital de documentos en Educación Matemática*. [En línea] 17 de agosto de 2012. <http://funes.uniandes.edu.co/>.
52. **Gómez, Pedro, Cañadas, María y Restrepo, Ángela.** Publicación y búsqueda de investigaciones en educación matemática: el aporte de Funes como repositorio digital de documentos. [En línea] <http://funes.uniandes.edu.co/1773/>.
53. **Instituto Tecnológico de Estudios Superiores de Occidente.** OAI y Metalib. El desarrollo de la biblioteca digital con CATIA. [En línea] mayo de 2005. [Citado el: 24 de agosto de 2012.] <http://quijote.biblio.iteso.mx/blog/GUAM/wp-content/uploads/2005/CatIA.ppt>.
54. **Universidad de las Américas Puebla.** Publications of the Interactive and Cooperative Technologies Lab. [En línea] [Citado el: 26 de agosto de 2012.] <http://ict.udlap.mx/pubs/index.html>.
55. **International Organization for Standardization.** ISO. [En línea] [Citado el: 23 de marzo de 2012.] <http://www.iso.org/>.
56. **O'Connor, Rory y Laporte, Claude.** Using ISO/IEC 29110 to harness process improvement in very small entities. [En línea] 2011. [Citado el: 12 de junio de 2011.] http://doras.dcu.ie/16723/1/10.1007_978-3-642-22206-1_20.pdf.

57. **Kendall, Kenneth y Kendall, Julie.** *Analisis y Diseño de Sistemas*. México : Pearson Educación, 2005.
58. **Bascón, Ernesto.** El patrón de diseño Modelo-Vista-Controlador (MVC) y su implementación en Java Swing. [En línea] 2004. [Citado el: 23 de agosto de 2012.] <http://ucbconocimiento.ucbca.edu.bo/index.php/ran/article/download/84/81>.
59. **Dickinson, Jon.** Grails 1.1 Web Application Development Packt Publishing Ltd. [En línea] 2009. [Citado el: 23 de agosto de 2012.] <http://books.google.com.mx/>.
60. **Brito, Nacho.** Manual de desarrollo web con Grails. [En línea] 2009. [Citado el: 24 de agosto de 2012.] <http://books.google.com.mx/>.
61. **SpringSource.** The Service Layer. *Grails*. [En línea] VMWare. [Citado el: 26 de agosto de 2012.] <http://grails.org/doc/latest/guide/services.html>.
62. **Apache Software Foundation.** Apache Tomcat. [En línea] [Citado el: 28 de agosto de 2012.] <http://tomcat.apache.org/>.
63. —. Apache HTTP Server. [En línea] [Citado el: 29 de agosto de 2012.] <http://httpd.apache.org/>.
64. **PostgreSQL Global Development Group.** PostgreSQL. [En línea] [Citado el: 28 de agosto de 2012.] <http://www.postgresql.org/>.
65. **World Wide Web Consortium.** HTML & CSS. W3C. [En línea] [Citado el: 26 de agosto de 2012.] <http://www.w3.org/standards/webdesign/htmlcss>.
66. **LongTail Video.** The State Of HTML5 Video. [En línea] abril de 2012. [Citado el: 26 de agosto de 2012.] <http://www.longtailvideo.com/html5>.
67. **WebKit Open Source Project.** WebKit Trac. [En línea] [Citado el: 26 de agosto de 2012.] <http://trac.webkit.org/>.
68. **Mozilla.** Ayuda de Firefox. [En línea] [Citado el: 27 de agosto de 2012.] <http://support.mozilla.org/>.
69. **World Wide Web Consortium.** W3Counter. [En línea] agosto de 2012. [Citado el: 26 de agosto de 2012.] <http://www.w3counter.com/globalstats.php>.
70. **Net Applications.** Net Applications Net Marketshare . [En línea] agosto de 2012. <http://www.netmarketshare.com/browser-market-share.aspx?qprid=1>.
71. **Unicode Resources.** UTF-8 and Unicode. [En línea] [Citado el: 27 de agosto de 2012.] <http://www.cl.cam.ac.uk/~mgk25/unicode.html#utf-8>.
72. **Montes, Daniel Y.** *Difusión de contenidos educativos mediante el uso de repositorios de objetos de aprendizaje y redes sociales*. México : s.n., 2012.
73. **Microsoft Corporation.** *Performance Testing Guidance for Web Applications*. s.l. : O'Reilly Media, 2010.
74. **Molyneaux, Ian.** *The Art of Application Performance Testing: Help for Programmers and Quality Assurance*. s.l. : O'Reilly Media, 2009.
75. **Barber, Scott.** Acceptable application response times vs. industry standard. *TechTarget*. [En línea] marzo de 2007. [Citado el: 25 de septiembre de 2012.] <http://searchsoftwarequality.techtarget.com/tip/Acceptable-application-response-times-vs-industry-standard>.
76. *Response time in man-computer conversational transactions.* **Miller, Robert.** San Francisco : ACM, 1968.
77. **Bickford, Peter.** Worth the Wait? [En línea] 1997. [Citado el: 27 de septiembre de 2012.] http://web.archive.org/web/20040202125921/http://developer.netscape.com/viewsource/bickford_wait.htm.

78. **Akamai, JupiterResearch.** Retail Web Site Performance: Consumer Reaction to a Poor Online Shopping Experience. [En línea] 2006. [Citado el: 27 de septiembre de 2012.] http://www.akamai.com/html/about/press/releases/2006/press_110606.html.
79. **Massachusetts Institute of Technology.** 2005 Program Evaluation Findings Report. [En línea] 5 de junio de 2006. [Citado el: 27 de septiembre de 2012.] http://ocw.mit.edu/ans7870/global/05_Prog_Eval_Report_Final.
80. **Czeiszperger, Michael.** Load Testing Basics: How Many Concurrent Users is Enough? *WebPerformance*. [En línea] febrero de 2011. [Citado el: 28 de septiembre de 2012.] <http://www.webperformance.com/load-testing/blog/2011/02/load-testing-basics-how-many-concurrent-users-is-enough/>.
81. **Apache Software Foundation.** Apache Lucene 4.0.0 Documentation. [En línea] [Citado el: 29 de septiembre de 2012.] http://lucene.apache.org/core/4_0_0/index.html.
82. —. Lucene-java Wiki. [En línea] [Citado el: 29 de septiembre de 2012.] <http://wiki.apache.org/lucene-java/FrontPage?action=show&redirect=FrontPageEN>.
83. **Krellenstein, Marc.** Full Text Search Engines vs. DBMS. [En línea] septiembre de 2009. [Citado el: 29 de septiembre de 2012.] <http://searchhub.org/dev/2009/09/02/full-text-search-engines-vs-dbms/>.
84. **International Institute for Software Testing.** Education-Based Software Test Certifications. [En línea] [Citado el: 10 de octubre de 2012.] <http://www.testinginstitute.com/>.
85. **International Software Testing Qualifications Group.** ISTQB. [En línea] [Citado el: 10 de octubre de 2012.] <http://www.istqb.org/>.
86. **American Software Testing Qualifications Board.** Software Testing Certification. [En línea] [Citado el: 10 de octubre de 2012.] <http://www.astqb.org/>.
87. **Chacon, Scott.** *Pro Git*. s.l. : Apress, 2009.
88. **Presman, Roger.** *Software Engineering: A Practitioner's Approach*. s.l. : McGrawHill, 2005.
89. **Tanenbaum, Andrew.** *Sistemas Operativos Modernos*. Ciudad de México : Pearson Educación, 2003.
90. **OECD.** Measuring Knowledge Management in the Business Sector: First Steps. [En línea] 2003. [Citado el: 20 de marzo de 2012.] http://www.keepeek.com/Digital-Asset-Management/oecd/science-and-technology/measuring-knowledge-management-in-the-business-sector_9789264100282-en.
91. **Flynn, Ida y McIver, Ann.** *Sistemas Operativos*. Ciudad de México : Cengage Learning Editores, 2001.