

Capítulo 1

Introducción

El libro ha sido una de las creaciones más valiosas del hombre. Desarrollado no sólo para aumentar la capacidad de su memoria, sino para establecer una forma de comunicación que le permitiera remontar las barreras del tiempo, el libro ha facilitado el intercambio de información útil y provechosa de cada época y de cada pueblo, logrando así la transmisión continua de conocimiento y el progreso del hombre a través de la historia.

Gracias a la existencia de bibliotecas alrededor del mundo que han resguardado material bibliográfico a través del tiempo, ha sido posible por medio del estudio y la investigación un proceso continuo de aprendizaje, ya que los libros representan una gran fuente de información y conocimiento. Sin embargo, mucho de este material enfrenta una problemática importante, principalmente en las áreas de preservación, difusión y consulta.

Por ejemplo, existen bibliotecas que cuentan con acervos especiales antiguos de gran valor no sólo por su contenido, sino por su trascendencia, estilo y conformación, que ya no pueden ser accesibles al público para su consulta debido a la fragilidad de sus materiales y a la necesidad que existe de protegerlos y conservarlos.

Debido a esta situación, la riqueza que pueden representar estos acervos en las diversas áreas del conocimiento no ha podido ser explotada, ya que no se cuenta con los mecanismos adecuados que permitan difundirla y consultarla además de aquellos que eviten la pérdida de dicho material.

La dificultad para consultar este tipo de libros es uno de los tantos inconvenientes que pueden representar las bibliotecas físicas, sumado a la falta de disponibilidad del material, el tiempo de préstamo el cual puede ser muy limitado, o las grandes distancias que

separan a los interesados. En contraste a los problemas que representa una biblioteca física, surgieron las bibliotecas digitales, las cuales cuentan con las herramientas y servicios que permiten dar acceso a grandes volúmenes de información de una manera fácil y práctica para los usuarios.

Sin embargo, a pesar de las ventajas que representa una biblioteca digital, muchos son los acervos que se encuentran aún en un formato físico enfrentando la problemática antes mencionada. Tal es el caso de la colección de la Biblioteca Franciscana, la cual motivó en gran medida este proyecto al contemplar precisamente el problema que representa el manejo de colecciones especiales.

Por otra parte, cabe recalcar que aunque la tesis se enfoca a trabajar sobre este acervo en específico, los resultados de la misma son de aplicación más general, ya que está desarrollada sobre un contexto que permite consultar información tanto de acervos bibliográficos de este tipo, como colecciones de diversa índole que presenten similitud en cuanto a la metodología empleada para la recuperación de información contenida en estas.

En términos generales, nuestro proyecto consistió en desarrollar un sistema para almacenar y consultar este tipo de libros en una biblioteca digital, contribuyendo a facilitar con ello el acceso, la difusión y la preservación de colecciones bibliográficas valiosas.

1.1 Bibliotecas digitales y U-DL-A

Las bibliotecas digitales han captado en los últimos años la atención de científicos e investigadores en diversas áreas (educación, historia, arte, ciencia, etc.) y han sido el resultado de los grandes avances tecnológicos y de la necesidad de manejar en forma eficiente grandes volúmenes de información.

Una biblioteca digital se puede definir como un espacio virtual que comprende colecciones digitales y proporciona de manera eficiente servicios como almacenamiento,

organización, recuperación, difusión y preservación de información así como acceso en línea a gran diversidad de material.

Las bibliotecas digitales representan grandes ventajas sobre las bibliotecas físicas en cuanto a disponibilidad de material, menor espacio físico de almacenamiento, mayor accesibilidad, gran diversidad así como las facilidades que brinda a los usuarios en cuanto a navegación, búsqueda, recuperación, y generación de la información [Lesk 1997].

Actualmente existe una gran diversidad de áreas en las que se han aplicado las bibliotecas digitales: en el ámbito médico, científico, artístico, noticioso, educativo, cultural, histórico, siendo esta una lista casi interminable.

La importancia que tiene para la sociedad el hecho de preservar material que documente eventos, hechos, e ideas, se ha visto reflejada precisamente en el surgimiento y desarrollo de ese tipo de tecnologías. En el aspecto histórico por ejemplo, la necesidad de preservar material bibliográfico es de vital importancia para el entendimiento de diversos acontecimientos, y principalmente como herencia cultural para futuras generaciones.

Actualmente, en la Universidad de las Américas, Puebla, está en marcha la iniciativa denominada U-DL-A (University Digital Libraries for All) desarrollada por el Laboratorio de Tecnologías Interactivas y Cooperativas (ICT) y la Biblioteca de la UDLA(ver <http://ict.udlap.mx>).

Esta iniciativa tiene la finalidad de apoyar la educación e investigación universitaria a nivel licenciatura y posgrado, bajo la cual se han integrado proyectos de bibliotecas digitales, explorando aspectos fundamentales en el desarrollo de estas.

1.2 Colecciones digitales dentro del contexto U-DL-A

Entre los acervos digitales que se están construyendo como parte de U-DL-A se encuentran la Colección de Tesis Digitales, la Correspondencia Presidencial de Porfirio

Díaz, La Colección de Publicaciones Universitarias, así como las Colecciones Especiales dentro de las cuales se encuentra contemplado el acervo bibliográfico de la Biblioteca Franciscana.

A continuación se ofrece una breve descripción de cada uno de los proyectos mencionados, y los avances que se han tenido en el desarrollo de estos, incluyendo el caso de la colección franciscana sobre la cual nos enfocaremos más adelante, debido a que merece una atención especial, ya que será punto de estudio para este proyecto de tesis.

1.2.1 Colección de Tesis Digitales

Considerando la importancia de las tesis no solo en el ámbito estudiantil, sino en el de investigación, debido a que son documentos formales con un análisis crítico y que presentan diagnósticos y soluciones a problemas reales, la UDLA-P inició el proyecto de Tesis Digitales como una forma de difundir el conocimiento, aprovechando las nuevas tecnologías de información para facilitar el acceso a este tipo de documentos, que presentaban problemas en cuanto a la falta de disponibilidad, los horarios de préstamo y la pérdida o mutilación del material.

Actualmente sólo se encuentra integrado un departamento académico, sin embargo, en un futuro se espera que todos los departamentos de la universidad se integren al nuevo formato de tesis digital (ver <http://mailweb.pue.udlap.mx/~tesis/>).

1.2.2 Correspondencia Presidencial de Porfirio Díaz

La correspondencia Presidencial de Porfirio Díaz (1876 a 1911) es una de las fuentes de información más valiosas de la historia moderna de México. La biblioteca de la Universidad de las Américas, Puebla alberga un total de 663, 843 manuscritos y telegramas así como cerca de 70,000 telegramas sobre los cuales ya se ha iniciado un proyecto de digitalización y decodificación.

Actualmente se tienen como resultados iniciales la versión digital de los telegramas generados en el año de 1910 (ver <http://digital.udlap.mx/porfirio.html>), a los cuales ya se puede tener acceso a través de herramientas de búsqueda y navegación.

1.2.3 Publicaciones universitarias

Las publicaciones universitarias como periódicos institucionales y estudiantiles así como otros documentos de interés histórico para la universidad, actualmente se encuentran contempladas dentro de un proyecto, en el que se pretende construir la colección tanto de publicaciones ya existentes así como de las actuales ya generadas en formato digital. Los elementos que conforman esta colección son principalmente libros, revistas y periódicos. Actualmente este proyecto está en proceso.

1.3 Colecciones especiales en las bibliotecas de la UDLA

Las colecciones especiales de las bibliotecas de la UDLA, se encuentran principalmente en su Sala Porfirio Díaz y en la recientemente incorporada Biblioteca Franciscana. Dado que la incorporación de esta biblioteca motivó el presente proyecto, se ofrecen aquí algunos de sus antecedentes y conformación.

La Universidad de las Américas, Puebla, en conjunto con la organización religiosa propietaria de bienes inmuebles y colecciones bibliográficas, llevan a cabo el proyecto denominado "Portal de Peregrinos", que consiste en la restauración del convento de San Gabriel, Cholula, con la finalidad de alojar la Biblioteca Franciscana y el Centro de Estudios Humanísticos Fray Bernardino de Sahagún (ver <http://biblio.pue.udlap.mx/portal/>).

En su afán de buscar siempre apoyar la difusión de conocimiento, el estudio y la investigación, la UDLA se ha propuesto desarrollar proyectos que permitan explotar la riqueza existente en colecciones bibliográficas de este tipo. Uno de estos proyectos es

precisamente el tema de esta tesis en la que se tiene como objetivo dar solución en parte a los problemas a los que se enfrenta el acervo, principalmente en las áreas de preservación, difusión y consulta.

1.3.1 Características e importancia de los libros

El acervo bibliográfico de la orden franciscana, esta formado por 25,000 volúmenes aproximadamente, con libros que datan del siglo XVI al XIX en idiomas como latín, español y francés entre otros. Los materiales que constituyen estos libros son principalmente cartón, pergamino y pieles, y la encuadernación que se conserva de algunos de ellos se caracteriza por su belleza, elegancia y su lujo.

Los libros que conforman la colección franciscana son considerados material muy valioso debido a que de una u otra manera señalan la cultura, la ciencia, el arte, la filosofía y teología en el tiempo y en el espacio humanos.

Muchos otros pueden ser los aspectos a tomar en cuenta para determinar el valor de los libros, su antigüedad, el hecho de haber sido escritos por algún autor célebre, porque el asunto de la obra es particularmente singular o ha sido tratado de manera singular, por su rareza o decoración, entre otros aspectos.

1.3.2 Procedencia de los libros de la orden franciscana

En esta biblioteca se tiene contemplado concentrar aquellos libros que tengan que ver con la orden franciscana en México. Este material bibliográfico proviene de varios conventos que citaremos a continuación:

En México:

Convento de Santa Úrsula

Convento de Coyoacán

En Puebla:

Convento de San Gabriel Cholula

Convento de San Cosme

Convento de San Antonio

Convento de Calpan

Convento de Puebla

En Veracruz:

Convento de Veracruz

Otros conventos

1.3.3 Situación actual de los libros

En la actualidad la mayor parte de los libros ya se encuentran inventariados para su asentamiento final en la Biblioteca Franciscana, sin embargo, aún falta por revisar cierta cantidad de material bibliográfico, ya que se desconoce la información necesaria que permita ubicarlos en el convento al que pertenecen, por lo que se encuentran clasificados en el grupo de otros conventos.

La información que se tiene recopilada con respecto a los libros que conforman este acervo es básicamente de tipo general como autores, títulos, países, editoriales, idiomas, tipos de material, tipos de documento, temas, editores, entre otros.

Parte de esta información puede ser recuperada a través de un sistema básico de consulta. Sin embargo, aún no se tiene disponible el formato digital del material bibliográfico, así como una aplicación que permita manipular dicha información y que haga posible efectuar búsquedas sobre el contenido de los libros, y la visualización del material para su análisis y estudio.

1.4 Resumen del problema

Como mencionamos en apartados anteriores, las tres áreas problema que enfrentan colecciones bibliográficas antiguas, y en este caso el acervo franciscano, son las de preservación, difusión y consulta.

La manipulación física sobre el material ya no es recomendable debido a que el estado de muchos de los libros es delicado por la antigüedad que estos presentan, y traería como consecuencia la pérdida parcial o total del mismo.

La consulta a los libros por lo tanto, se torna difícil, principalmente para usuarios interesados en este tipo de libros que se encuentran a grandes distancias de la biblioteca que aloja el material bibliográfico. Tomemos en consideración además que la información actual con la que se cuenta sobre estos libros es de tipo general no de contenido, lo que limita en mucho el acceso a la información que pudiera ser de interés e importancia para usuarios de este tipo de colecciones.

Finalmente, la falta de difusión constituye otro problema, esto principalmente por que no se cuenta con un formato adecuado que lo haga posible, y que permita no solo dar acceso a la información de manera local sino remota.

1.5 Mecanismos de preservación y consulta

Para hacer frente a esta problemática, se planteó como objetivo de este proyecto investigar métodos existentes para almacenamiento y consulta de textos digitalizados, así como el desarrollo e implementación de un sistema que permitiera a los usuarios de este tipo de acervos tener acceso al material sin necesidad de manipularlo físicamente. Para facilitar la difusión de la información contenida en este tipo de colecciones se planteó también como objetivo que el sistema desarrollado fuera accesible vía web.

El sistema resultante denominado CITexten alusión a su funcionalidad en cuanto a la Consulta de *Imágenes Textuales*, facilita la consulta y recuperación de información,

permitiendo al usuario visualizar la reproducción digital del original, que finalmente puede navegar como si lo tuviera físicamente en sus manos.

Siendo una parte importante la información de tipo general como autores, títulos, temas, clasificaciones, el software cuenta con un sistema de búsqueda básico, que también permite recuperar información de esta índole.

1.6 Estructura del documento

El documento está organizado de la siguiente manera. El capítulo 2 se enfoca a revisar trabajo previo relacionado en cuanto a preservación y metodologías de búsqueda. El capítulo 3 muestra aspectos de diseño del sistema así como el modelo de datos, software comercial de reconocimiento utilizado, análisis de algoritmos a utilizar para la búsqueda de información, interfaces con el usuario y funcionamiento del sistema. En el capítulo 4 se detallan aspectos de implementación. En el capítulo 5 se muestra un recorrido por el sistema implementado y finalmente en los capítulos 6 y 7 se consideran las pruebas de los algoritmos propuestos, resultados y conclusiones.