

# Capítulo I: Introducción

Es fin de éste capítulo dar una descripción general de éste trabajo de tesis. Se mencionan los diferentes objetivos, limitaciones y alcances sobre los cuales se desarrolló la tesis y se describe brevemente la estructura y el funcionamiento general del sistema desarrollado. También se describe brevemente el contenido de cada uno de los capítulos contenidos dentro de éste documento de tesis.

## 1.1 Introducción y Descripción del Problema

Las ciencias de la computación, es un área de las ciencias la cual cambia constantemente. La gran preocupación es que los avances tecnológicos, especialmente en esta área, son tan constantes y tan rápidos que es muy difícil mantenerse al día y desarrollar técnicas y herramientas las cuales puedan realizar las tareas necesarias para que las cosas funcionen eficientemente. Actualmente nos encontramos con el dilema de la búsqueda y recuperación eficiente de información de cualquier tipo. La tarea de buscar información que nos pueda ayudar es muy tardada y difícilmente los resultados que obtenemos son los deseados. Lo que necesitamos es desarrollar técnicas y herramientas las cuales nos ayuden a realizar búsquedas rápidas y eficientes y que podamos realizarlas hacia cualquier tipo de información o datos existentes.

Este trabajo de tesis abarcó varias áreas diferentes de estudio las cuales se complementan para formar el sistema que se desarrolló a lo largo de la tesis. Una de las áreas de estudio, es la eficiente búsqueda y recuperación de información. En esta tesis se estuvo trabajando con la realización de búsquedas sobre datos no estructurados, en específico sobre archivos de audio llamados MP3. ¿Cómo realizar búsquedas sobre datos no estructurados, cómo realizar búsquedas sobre archivos de audio? ¿Cómo recuperar información de manera rápida y eficiente? ¿Cómo presentarle al usuario los resultados de manera ordenada del archivo más relevante al menos relevante? Estas son algunas de las preguntas que éste trabajo trato de responder. Se desarrolló una herramienta actual y funcional con la cual el usuario pueda realizar búsquedas de archivos de audio, en específico MP3, obteniendo resultados confiables y útiles.

Actualmente uno de los temas tecnológicos de controversia más grande es el tráfico de archivos MP3 por Internet. Hoy en día la industria musical se siente amenazada por el tráfico de archivos de audio MP3 que existe por toda la red; decenas de programas tipo *P2P* (programas que comunican dos computadoras directamente permitiendo el intercambio libre de cualquier tipo de archivo) se encargan de transferir miles de canciones por día. El problema de estos servicios es que ninguno de ellos se interesa por verificar si los archivos tienen derechos de autor, o si el nombre del archivo junto con todas sus etiquetas corresponden realmente a lo que el archivo contiene, o si el archivo fue creado correctamente y si este está en buen estado, completo, sin virus, etc.

Otra de las áreas de estudio con las que se trabajó, es el *web dinámico*, la administración de un portal en Internet y el manejo de diferentes tipos de usuarios dentro del mismo sistema. Muchos de estos problemas se relacionan con la seguridad y con la coherencia de la información. Este no es un sistema dentro del cual la seguridad sea tan crítica, como el de un banco por ejemplo, ya que no haremos transacciones de dinero ni manejaremos información vital de personas como números de tarjetas de crédito. Pero si fue de interés construir un sitio confiable donde se mantengan bien marcados los límites de acceso de cada uno de los diferentes tipos de usuarios y donde la información de los usuarios y de los archivos se guarde con total seguridad y coherencia.

El sistema también se construyó de tal manera que cualquiera de los usuarios lo pueda utilizar fácilmente y sin ninguna complicación, y que la persona asignada para administrar dicho portal, lo pueda hacer con total *transparencia*; sin ningún conocimiento avanzado de computación y sin que tenga que escribir ni una línea de código.

## 1.2 Objetivos Generales

- Desarrollar un portal de Internet con el cual podamos buscar, obtener y agregar archivos de audio MP3.
- Desarrollar un sistema con el cual se pueda administrar completamente dicho portal con total transparencia para todos los usuarios.

### 1.3 Objetivos Específicos

- Desarrollar todas las aplicaciones necesarias para que el administrador, el proveedor y el usuario final, puedan utilizar el sistema con total transparencia.
- Optimizar el sistema de consultas combinando varias técnicas de recuperación de información y de realización de búsquedas.
- Desarrollar las herramientas del proveedor: subir archivos, modificar su perfil, actualizar sus datos, actualizar su información y sus *links*.
- Desarrollar las herramientas de los usuarios del sistema: el buscador, el *raking* de los resultados y la capacidad de bajar los archivos.
- Desarrollar las herramientas del administrador: altas y bajas de proveedores, actualizar la página principal con nuevos artículos e imágenes.

### 1.4 Alcances

- El sistema, si es que se llega a publicar en Internet, podrá ser accedido por cualquier persona con acceso a Internet; solo podrá buscar y bajar archivos, no podrá agregar archivos a la colección.
- El administrador y todos los proveedores gozarán de un perfil el cual contendrá toda su información, e interfaces personalizadas.
- Las búsquedas se realizarán sobre los *metadatos* contenidos en los índices y no sobre los archivos ni la base de datos relacional.
- Para la tesis se utilizaron archivos MP3 obtenidos de los discos originales.
- El sistema maneja solamente archivos de audio MP3.
- El sistema maneja tres tipos de usuarios distintos con servicios personalizados.

## 1.5 Limitaciones

- Consideramos que solo habrá un administrador el cual será la misma persona que actualice la página con sus artículos nuevos, actualice las fotografías y realice todas las tareas administrativas del sitio. No podrá dar de alta a otro administrador, reportero, fotógrafo, editor, etc.
- Las búsquedas no se realizan sobre los archivos, ni sobre sus cadenas o patrones de cadenas binarias.
- La información de las transacciones realizadas por el sistema, no se encuentra *cifrada*.
- No se realiza ningún tipo de transacción monetaria, ni de información bancaria de ningún tipo.
- Los usuarios finales del sistema no tienen una cuenta personal, ni un perfil, ni interfases personalizadas; a todos se les manejará por igual.
- El sistema no maneja archivos de video, imágenes u otro tipo de archivo de audio que no sea MP3.
- No se busca virus en los archivos agregados a la colección por parte de los proveedores.
- El sistema no se subirá a Internet, o sea, no estará accesible al público. El sistema sí funcionará pero no se publicará, de esta manera se podrán utilizar los archivos de audio obtenidos de los discos originales, comprados previamente, para uso personal sin violar ninguna ley de derechos de autor.
- No se contactó ni a disqueras ni a artistas para que nos proporcionen archivos ni ningún tipo de información.

## 1.6 Descripción del Documento de Tesis

- **Capítulo I** – (Introducción) Descripción general del sistema y del documento, aquí se encuentran los objetivos, las limitaciones y los alcances de dicho trabajo de tesis.

- **Capítulo II** – (Marco Teórico) Se describe el marco teórico sobre el cual éste trabajo de tesis se encuentra fundamentado. Este capítulo describe todas las bases teóricas de éste trabajo de tesis.
- **Capítulo III** – (Descripción del Sistema) Descripción detallada del sistema. En éste capítulo describimos a detalle el análisis de requerimientos y el análisis del diseño de nuestro sistema.
- **Capítulo IV** – (Pruebas y Conclusiones) En este capítulo describimos las distintas pruebas realizadas a nuestro sistema para medir su robustez, así como los resultados obtenidos de dichas pruebas. En éste capítulo también se describen las distintas conclusiones a las que se llegaron durante el desarrollo de éste trabajo de tesis, y finalmente se describe el trabajo a futuro por realizar, es decir, todo el trabajo que un futuro se pueda continuar dentro de ésta línea de investigación.
- **Apéndice A** – (Glosario) En éste apéndice se definen los términos más importantes utilizados en éste documento de tesis.
- **Apéndice B** – (Tablas y Figuras) Este apéndice contienen las distintas tablas y figuras a las cuales se hace referencia dentro de éste documento de tesis.
- **Bibliografía** – Describe detalladamente toda la bibliografía utilizada durante todo el desarrollo de éste trabajo de tesis.