

ARTICULOS

CABSYS BASE DE DATOS BIBLIOGRÁFICA SOBRE RELIGIÓN EN COSTA RICA, NORMALIZACIÓN Y VALORACIÓN DE CONTENIDOS

*Aracelly Ugalde Víquez**

Resumen

Esta base de datos contiene 30738 registros referenciales de la prensa religiosa costarricense, en formato CEPAL, elaborados utilizando el programa Microsis, en la Escuela Ecuménica de la Religión de la Universidad Nacional.

Se presenta la experiencia y los procedimientos utilizados para la normalización de la base de datos Cabsys, desde la identificación y verificación de los artículos ingresados desde 1986, directamente de las fuentes originales: los periódicos religiosos, la complementación de los títulos, la elaboración de un archivo de autoridad de autor, la corroboración de los datos referenciales de ubicación; la normalización de los descriptores y la elaboración de un archivo de autoridad de materia.

Se presentan cuadros ilustrativos sobre los artículos incluidos, por ejemplo: distribución de trabajos por autor, distribución por género, distribución por año de publicación, temas más representativos según descriptores asignados y correlación entre estos y las palabras del título de los artículos.

Introducción

El control bibliográfico de la documentación religiosa es un requerimiento valioso para las investigaciones y se fundamenta en dos aspectos: el primero es facilitar el acceso al conocimiento sobre la cultura religiosa, su influencia y el legado que representa en Costa Rica. El segundo es ayudar al usuario a localizar la documentación y orientarlo en su búsqueda, en las diferentes localizaciones físicas de los

* Académica de la Escuela de Bibliotecología, Documentación e Información de la Universidad Nacional

** Revisado por MILS. Lucía Chacón Alvarado, Vicedecana, Facultad de Filosofía y Letras, Universidad Nacional

documentos, debido a que los documentos están en la Curia y los prestan de uno en uno, el servicio no es en horario continuo, ni diario, no disponen de una base de datos y el papel periódico utilizado esta altamente deteriorado, a punto de destruirse en el caso de los documentos más antiguos.

La Escuela Ecuménica de la Religión cuenta con la herramienta Cabsys, que es el primer intento de registrar información en esta área. Esta base de datos fue elaborada por personal asignado a este proyecto, con la colaboración profesores y estudiantes de la Escuela de Bibliotecología, Documentación e Información.

Cabsys es una amplia base de datos referencial con 30738 citas bibliográficas de artículos aparecidos en periódicos costarricenses sobre religión. Este esfuerzo es un aporte profesional de alto costo, una herramienta de consulta documental valiosa por su numeroso contenido referencial e histórico de artículos seleccionados por profesores de la Escuela Ecuménica de la Religión con base en criterios como: importancia histórica, profundidad, representatividad, perfil de interés, autor, vigencia, relevancia, etc., tomados de la prensa costarricense religiosa desde 1890 hasta 1991.

Los registros originales de la base de datos se almacenaban originalmente en cinta magnética (respaldo electrónico de información que se utilizó en los años 80's) y estas se dañaron, por lo que fue necesario la redigitación de la información a partir de una impresión, lo que pudo originar variaciones, omisiones y errores de digitación. La base de datos tenía muchos errores de normalización, digitación, ortografía y la abreviación de palabras para reducir los espacios de memoria en el computador.

Estos errores generaban problemas de precisión tales como: ruido, basura, silencios y dificultad en el uso de las referencias bibliográficas e imposibilitaban la valoración de la información particularmente dificultaba la realización de otros estudios por ejemplo: temas con más representatividad, periódicos analizados, autores con más artículos.

Este problema se solucionó aprovechando la oportunidad que se presentó de corregirla y la voluntad de las autoridades de la Escuela Ecuménica de la Religión para ponerla al servicio de los investigadores en esa área del conocimiento.

El objetivo final de este trabajo fue contribuir con el control bibliográfico de la documentación en el área de la religión, y compartir la experiencia aprendida en este proceso de corrección deja lecciones importantes para futuros trabajos similares con bases de datos que pueden ser útiles para estudiantes de la carrera de bibliotecología, documentación e información, para investigadores y profesionales dedicados a la depuración de bases de datos bibliográficas.

Metodología

Se elaboró un diagnóstico de la situación en relación con las referencias bibliográficas existentes en la Base de Datos y la disponibilidad de recursos para su depuración, además, se identificó el entorno de trabajo y la factibilidad de completar la depuración de la Base de Datos Cabsys. El estudio consistió en analizar las características de la información contenida en los registros de la base de datos, se elaboró una versión preliminar con la información y cinco índices: autor, título, materia, título del periódico y cronológico. Durante este ejercicio se determinaron los errores y fallas que se requería corregir, tales como:

Autores: Entradas sin normalizar, generalmente apellidos y nombres abreviados, incompletos, con errores de digitación, puntuación y ortográficos. El primer índice de autor generó un listado con 3680 entradas correspondientes a 12824 ligas a autores. El listado final de autores normalizado generó un índice de 2006 entradas y 12228 ligas a autores.

Títulos: Los títulos variaban mucho del original por el uso de abreviaturas usadas indiscriminadamente debido al reducido espacio disponible en la base de datos original, se revisaron y corrigieron los títulos de los 30738 registros.

Nombre del periódico: Los nombres títulos de las revistas estaban registrados sin normalización, por esta razón al generar un índice, este contenía 52 entradas en lugar de las siete entradas correspondientes al número de periódicos incluidos.

Descriptor: Los términos utilizados como descriptores no estaban normalizados, se generó un primer índice que presentó 2943 entradas de descriptores con varias entradas incorrectas. Al concluir el trabajo se generó un índice con 861 entradas de descriptores normalizados.

A partir de los resultados del diagnóstico se estructuró una propuesta estratégica para corregir los errores y se determinaron las tácticas para realizar el trabajo necesario de depuración de la base de datos. En el diagnóstico se determinó además, la factibilidad operacional, técnica y presupuestaria necesaria ejecutar la propuesta que fue aprobada por el coordinador del proyecto.

La propuesta consistió en iniciar la depuración de los registros directamente con las fuentes de información: las publicaciones periódicas, la sede donde se almacenan es la Curia metropolitana. Se designaron 11 estudiantes de informática del Instituto Profesional de Educación Comunitaria (IPEC) y se les distribuyó el material impreso para que identificaran los errores. Realizar este trabajo tubo una duración de dos meses.

La normalización de las entradas de los autores se realizó con la colaboración de Roxana Castro y Fernando Villegas, quienes revisaron el índice de autor y determinaron y construyeron las entradas correctas.

Para la normalización de las entradas de materia, se utilizaron los siguientes instrumentos: Tesauro de Religión y Lista de encabezamientos de materia para teología.

Las estudiantes de Bibliotecología se encargaron de depurar la información en la base de datos, la labor consistió en verificar la información, corregir y normalizar los errores: abreviaciones, omisiones y entradas incorrectas y diferenciadas. Al terminar las jornadas de trabajo se realizaba un respaldo de la base de datos en el disco duro, en discos de 3 ¼ y en disco zip.

El software utilizado fue Microisis versión 307, MS. Office 97 y Winisis versión 3. El hardware utilizado para trabajar fue un computador personal Pentium con unidad zip. El trabajo se llevó a cabo en las instalaciones de la Escuela Ecuménica de la Religión donde se dispuso del espacio físico, mobiliario y materiales de oficina.

El proceso duró 10 meses y contó con la participación de 18 personas, profesionales y estudiantes.

Implementación de la propuesta Depuración de Cabsys: guía del proceso

Procesos	Acciones
Normalización de materia	<ol style="list-style-type: none"> 1. Generación del archivo invertido de la base de datos e impresión del índice de descriptores, utilizando el archivo invertido para generar únicamente el campo de descriptores en la tabla de selección de campos (fst). 2. Identificación de los descriptores correctos y marcación de los mismos, revisando tesauros en la temática religiosa (ver bibliografía consultada) y consultando a los facilitadores. 3. Se elaboró un índice temático para depurar los descriptores con base en los tesauros. Al ser tan grande la base de datos y poca la capacidad del hardware disponible, cada corrección global consumía 1 minuto y 30 segundos, muy lento. 4. Muchos descriptores eran entradas de materia y se dividieron para convertirlas en descriptores según fuera necesario. 5. Gran cantidad de descriptores estaban abreviados y fue necesario cambiarlos por la palabra completa, por ejemplo JC en Jesucristo, Pres. por Presbítero y AT por (A.T.) 6. Corrección de los descriptores con errores por medio de cambios globales cuando eran más de diez postings (cantidad de registros de la base de datos que contienen un descriptor). 7. Edición de los registros que contenían errores en los descriptores con menos de diez postings. 8. Volver al punto 1 para chequear correcciones y errores en listados impresos con depuraciones. 9. La corrección de los descriptores se realizó en un cuarto de tiempo durante 9 meses.
Normalización de autor	<ol style="list-style-type: none"> 1. Generación del archivo invertido de la base de datos e impresión del índice de autor, utilizando el archivo invertido para generar únicamente el campo de autor en la tabla de selección de campos (fst). 2. Identificación de los autores y sus entradas correctas, marcación de los mismos. En esta etapa se contó con la colaboración de la Licda. Roxana Castro, profesora egresada de la Escuela de Bibliotecología, Documentación e Información en la determinación de las entradas de autor correctas. 3. Corrección de los autores con errores en las entradas por medio de cambios globales cuando eran más de diez postings. 4. Edición de los registros que contenían errores con menos de diez postings. 5. Volver al punto 1 para chequear correcciones y errores en listados impresos con depuraciones. 6. La corrección de las entradas de autor se realizó en un cuarto de tiempo durante ocho meses.
Procesos	Acciones

Identificación y verificación de los artículos referenciados	<ul style="list-style-type: none"> - Se generó un listado impreso de toda la base de datos ordenada por el nombre de la publicación periódica, dentro de esta clave, ordenada por año, volumen, número y por páginas del artículo. - Esta impresión contenía 3200 hojas con un promedio de 16 registros por hoja. - Se contrató a once estudiantes de informática y se les entregó un promedio de 200 hojas a cada uno para que visitaran la Curia Metropolitana y la Biblioteca Nacional y corroborarán la información impresa con la información en la fuente. - Los estudiantes trabajaron durante un mes y medio (35 horas) corrigiendo títulos, verificando autores, periódicos, fechas y páginas en la fuente. - Se detectaron muchos artículos duplicados, errores y abreviaciones, las correcciones se anotaron en los registros impresos. - Las correcciones y observaciones se anotaron en las impresiones y documentos adjuntos. - La bibliotecóloga encargada de depurar títulos trabajó un cuarto de tiempo durante tres meses en la corrección de los registros en la base de datos. - Se eliminaron 1236 artículos duplicados. - Se imprimió un índice de título y se evaluaron los resultados.
Títulos y descripción física	<ul style="list-style-type: none"> - Se realizó una revisión intensiva de la base de datos en pantalla, registro por registro para detectar y corregir errores. - Se hizo un chequeo con tres revistas del Eco Católico y cinco de la Hojita parroquial, verificando artículo por artículo en la base de datos.
Periódicos	<p>El proceso realizado fue el mismo que se realizó con el autor y los descriptores, se imprimió el listado de periódicos (adjunto), se identificaron las entradas correctas y se procedió a corregir las otras.</p>
Listados y respaldos	<ul style="list-style-type: none"> - Se imprimió toda la base de datos en orden del periódico y en orden cronológico por fecha. - Se imprimieron 10 listados de autor, 10 de descriptores, tres de periódicos.
Coordinación y orientación	<ul style="list-style-type: none"> - Generación de archivos para validad entradas, - Revisión de los impresos, - Verificación y elaboración de respaldos y acciones en relación con el tiempo asignado: un año. - Capacitación y orientación durante todo el proceso. - Elaboración de informes parciales y final. - Cada respaldo elaborado consumía 5:00 minutos en los diferentes discos de almacenamiento.

Resultados y valoración del contenido de la Base de datos Cabsys

La base de datos Cabsys depurada permite realizar valoraciones de contenido muy interesantes para conocer la distribución histórica de los artículos registrados, los autores en relación con la cantidad de artículos publicados por ellos, las temáticas abarcadas según los descriptores utilizados y según las palabras contenidas en el título.

Por ejemplo, los temas más representativos son: iglesia, clero y eclesiología con 6746 artículos, Biblia con 1773, misiones 1497, fe 1497, Jesucristo 1338, teología 1252, etc. De 2006 autores, solo 58 son mujeres.

La distribución histórica de los artículos por décadas, se ilustra en la figura 1:

Cantidad de artículos que contiene la Base de Datos Cabsys por década, desde 1880 a 1999

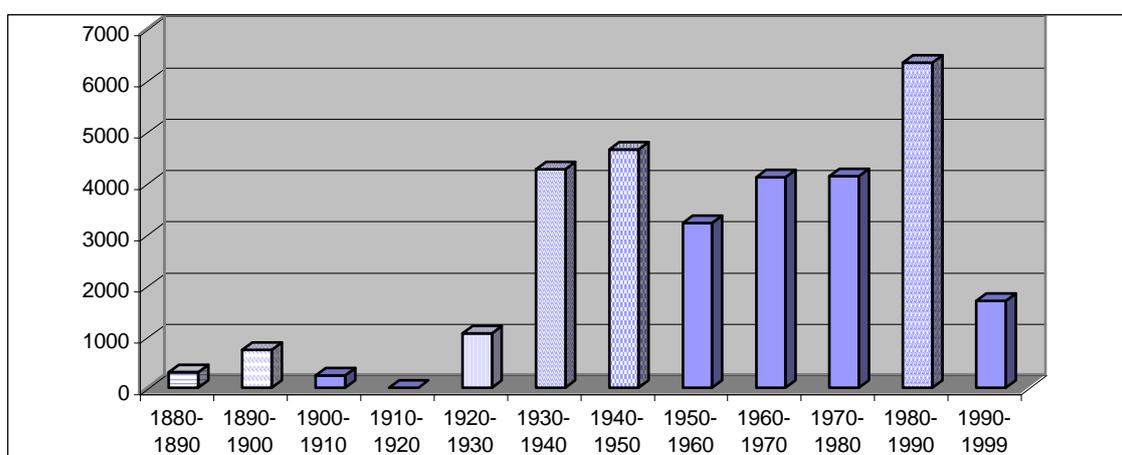


Figura # 1

Solamente en la década de 1910 no hay ninguna referencia bibliográfica.

Autores con más artículos

Figura #2

Conclusiones

<u>Autor</u>	<u>Articulos</u>	<u>Autor</u>	<u>Articulos</u>
Borge C., Carlos, Presbítero	549	Vargas C., Juan	126
Juan Pablo II, Papa	494	Castro Jiménez, Rafael Oton,	124
Mendoza, Juan Luis, Presbítero	485	Arzobispo	
Arrieta, Roman, Arzobispo	437	Ivanhoe, Seud.	119
Guillen Quirós, Rafael Maria	410	Castillo Vega, Abel, Presbítero	113
Arrieta, Victor Manuel, Presbítero	406	Rodríguez, Carlos, Arzobispo	111
Troyo, Antonio, Obispo Auxilia	242	Alfaro Odio, Carlos Joaquín, Presbítero	93
Bolanos, Enrique, Obispo	175	Arias, Elias, Presbítero	92
Mata Oreamuno, Alberto, Presbítero	170	Trejos, Oscar José, Presbítero	92
Munoz, Gonzalo, Presbítero	166	Quirós C., José Alberto, Presbítero	90
Odio, Ruben, Arzobispo de San	162	Pacelli, Eugenio, Cardenal	88
Castro Herrera, Alicia, Sor	157	Solís, Javier, Presbítero	85
Alfaro, Armando, Presbítero	155	Bolanos, Claudio, Presbítero	80
Sanabria, Victor Manuel, Arzobispo	155	Campos Sandí, Jorge, Presbítero	78
Riano, Eloy, Presbítero	145	Pauels, José, Presbítero	72
Meneses B., Carlos, Presbítero	141	Sandí, Walter, Presbítero	72
		Trejos, Ignacio, Obispo	67

El proyecto permitió capacitar a tres estudiantes de bibliotecología en la ejecución de procesos de depuración de bases de datos y en la realización de proyectos interdisciplinarios con el enriquecimiento que estos representan para la formación profesional

Además, se cumplió con el objetivo de contribuir con el control bibliográfico en el área de la religión en Costa Rica, pero es el primer paso, debe ser completado con la información que falta en cada año y con el texto completo, para publicarlo en Internet y que pueda utilizarse por los investigadores e interesados a nivel nacional e internacional.

Por esta razón, se recomienda continuar este esfuerzo con el proyecto "Escribiendo a Jesús", que consistirá en el rescate del patrimonio documental de la religión católica en Costa Rica mediante la digitalización, el procesamiento y la sistematización de la información escrita en la prensa costarricense y la conversión de estos recursos en una biblioteca electrónica a la que los investigadores accedan por medio de un portal publicado en la red Internet.

Agradecimiento

Especialmente la Profesora Licda. Roxana Castro Alpizar, profesora egresada y exdirectora de la Escuela de Bibliotecología, Documentación e Información y del Lic. Fernando Villegas Alfaro, profesor, exsubdirector de la Escuela Ecuménica de la Religión, quienes revisaron el índice de autor y determinaron y construyeron las entradas correctas para la normalización de las entradas de los autores se realizó.

Otros colaboradores muy importantes del proceso fueron:

- Dr. Miguel Picado Gatengs, director de la Escuela Ecuménica de la Religión, coordinador del proyecto.
- Lic. Fernando Villegas Alfaro, subdirector Escuela Ecuménica de la Religión.

Estudiantes:

Escuela de Bibliotecología, Documentación e Información:

- Lilliam Núñez
- Grettel Murillo
- Shirlene Ugalde

11 Estudiantes de Informática del Instituto Profesional de Educación Comunitaria.

Literatura Consultada

Pérez, Alvaro. Lista de encabezamientos de materia para teología. San José: Red Latinoamericana de Información Teológica, 1998.

Sequeira, Deyanira y Picado, Miguel. Tesauro de religión. Heredia: Universidad Nacional, 1980.

Micro CDS/ISIS: versión 3.7 y versión winisis. Paris: UNESCO, 1999.