12
29,

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE ESTUDIOS SUPERIORES "ZARAGOZA"

DICCIONARIO COMPUTARIZADO DE NOMBRES DE TAXONES DE PLANTAS VASCULARES Y VERTEBRADOS PARA BASES DE DATOS DE COLECCIONES BIOLOGICAS

T E S I S

QUE PARA OBTENER EL TITULO DE

B I O L O G O

P R E S E N T A N

TANIA ESCALANTE ESPINOSA

JOSE GERARDO RODRIGUEZ TAPIA

DIRECTOR: DAVID NAHUM ESPINOSA ORGANISTA.

FACULTAD DE ESTUDIOS SUPERIORES "ZARAGOZA", UNAM

CO-DIRECTOR: VERONICA AGUILAR SIERRA.

COMISION NACIONAL PARA EL CONOCIMIENTO Y USO DE LA BIODIVERSIDAD (CONABIO)

U.N.A.M.
FES
ZARAGOZA

LO HUMANO EJE DE NUESTRA REFLEXION

SEPTIEMBRE DE 1998

265944

TESIS CON
FALLA DE ORIGEN

Universidad Nacional Autónoma de México
UNAM
Biblioteca Central
Dirección General de Bibliotecas de la UNAM

UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Primeramente queremos hacer constar que parte del desarrollo de este trabajo se llevó a cabo con el apoyo del proyecto IN207995 de la DGAPA "Áreas de endemismo del trópico mexicano", del cual fungió como responsable el Dr. Jorge Llorente Bousquets y co-responsable el M. en C. David N. Espinosa Organista, en el cual participamos como becarios.

Deseamos expresar nuestro reconocimiento a todo el personal de la CONABIO, principalmente quienes laboran actualmente en la Subdirección de Inventarios Bióticos (Biól. Patricia Koleff Osorio y Biól. Verónica Aguilar Sierra), en la Subdirección de Desarrollo y en la Dirección de Sistemas (Lic. Pablo Ortuño). También agradecemos a Julia Moreno, Miguel Ángel Granados, Víctor Amezcua y Rafael Caballero por sus sugerencias y enseñanzas.

Igualmente agradecemos las valiosas opiniones y sugerencias de cada uno de los especialistas que contribuyeron en el aporte de ideas para la elección de los diferentes sistemas de clasificación.

Además, de manera muy especial a David N. Espinosa Organista y Rafael Mayorga Saucedo.

Asimismo durante el desarrollo de este trabajo han colaborado con nosotros muchas personas quienes con sus ideas y comentarios hicieron posible que se terminara con éxito. Finalmente ofrecemos de antemano una disculpa a quienes no mencionamos, más sin embargo ellos saben lo valiosas que fueron sus aportaciones.

DEDICATORIA

"Todos para uno y uno para todos"
Les Trois Mousquetaires,
Alexandre Dumas

Por su cariño y comprensión en cada instante,
porque siempre estaremos juntas
Sheila y Erika

Por su apoyo incesante,
con mi más grande respeto y admiración
Luis Felipe

Por haberme dado el privilegio de vivir,
por quien estoy aquí
Silvia

Por ser la persona con la que quiero
compartir los momentos más importantes de mi vida
Gerardo

Tania

DEDICATORIA

A Carlos:
Mi ejemplo para luchar por la vida

A Rocío:
Mi sonrisa en los momentos difíciles

A Andrea:
Mi silencio alentador

A Conchita:
El más grande ejemplo de fuerza y entereza

A mis padres:
A mis padres, a ellos...todo

A Tania:
Mi motivo de lucha y toda, toda mi felicidad

Gerardo

CONTENIDO

RESUMEN i

I. INTRODUCCIÓN 1

1.1 Colecciones biológicas científicas 1

1.2 Los sistemas de clasificación nomenclaturales 3

1.2.1 Sistemas de clasificación para plantas vasculares 4

1.2.2 Sistemas de clasificación para vertebrados 6

1.2.2.1 Sistemas de clasificación para mamíferos 6

1.2.2.2 Sistemas de clasificación para peces 7

1.2.2.3 Sistemas de clasificación para reptiles y anfibios 8

1.2.2.4 Sistemas de clasificación para aves 8

1.3 La computarización de colecciones biológicas 9

1.3.1 Bases de datos relacionales 10

1.3.2. Software para bases de datos 11

1.3.3 Modelos de bases de datos relacionales 14

1.3.4 Implicaciones y limitantes en el manejo y uso de bases de datos curatoriales 16

1.4 Bases de datos para consultar y administrar colecciones 18

1.4.1 Esfuerzos para computarizar colecciones 20

1.4.2 El modelo de datos BIOTICA 23

1.4.3 Diccionarios de autoridades y catálogos controlados 24

II. OBJETIVOS 29

III. MÉTODO 30

IV. RESULTADOS 34

4.1 *Sistemas de clasificación para plantas vasculares y vertebrados* 34
4.2 Descripción del modelo de la base de datos 36
4.3 Descripción de los objetos de la base de datos 36
4.3.1 Entidades 37
4.3.2 Consultas 40
4.3.3 Formularios 42
4.3.4 Reportes 44
4.4 Descripción de los nombres de taxones 46

V. ANÁLISIS DE RESULTADOS 50

5.1 *Colecciones biológicas y sistemas de clasificación* 50
5.2 Modelo Entidad-Relación: tratamiento de los datos nomenclaturales 52
5.3 Bases de datos nomenclaturales 59
5.4 Implementación del modelo de datos 60
5.4.1 Sección nomenclatural de BIOTICA 60
5.4.2 BIOTICA y otros modelos de datos 67

VI. CONCLUSIONES 69

VII. LITERATURA CITADA 71

ANEXO 1 78

GLOSARIO 81

ÍNDICE DE FIGURAS

Figura 1. Tipos de relaciones entre entidades. 15

Figura 2. Entidades básicas de una base de datos relacional curatorial cuyo centro es la información del ejemplar. El objeto de estudio de este trabajo es la entidad nomenclatural. 19

Figura 3. Tablas básicas del modelo de datos BIOTICA (CONABIO, 1998). 24

Figura 4. Proceso de control de calidad llevado a cabo en los datos que son incorporados al SNIB (Soberón and Koleff, 1998). 26

Figura 5. Catálogos desarrollados para llevar a cabo el control de calidad de las bases de datos del SNIB. 27

Figura 6: Modelo de datos BIOTICA. 28

Figura 7. Esquema del árbol de categorías ascendentes obligatorias y nombres incluido en la entidad con relaciones unarias. 31

Figura 8. Esquema del árbol de categorías ascendentes y nombres incluido en la tabla recursiva. 31

Figura 9. Representación esquemática del método. 33

Figura 10. Relaciones de las entidades de la base de datos. 36

Figura 11. Vista de la ventana de Access que muestra las entidades o tablas de la base de datos. 37

Figura 12. Ejemplo de consulta para obtener la clasificación completa de anfibios con las categorías obligatorias. 41

Figura 13. Ejemplo de consulta para actualizar registros en la clasificación de mamíferos. 42

Figura 14. Ejemplo de formulario para ingresar autores nuevos. 43

Figura 15. Ejemplo de formulario para captura de nombres y ascendentes de plantas vasculares. 43

Figura 16. Ejemplo de formulario para revisar y depurar la clasificación de aves. 44

Figura 17. Ejemplo del diseño de un reporte para el sistema de clasificación de anfibios. 45
Figura 18. Ejemplo de la presentación preliminar de un reporte para el sistema de clasificación de anfibios. 45
Figura 19. Cambios en la circunscripción de un taxón, según Berendsohn (1998). 52
Figura 20. Modelo de datos donde cada entidad corresponde a una jerarquía taxonómica. 53
Figura 21. Modificación al modelo de datos donde se ingresa una categoría intermedia. 53
Figura 22. Esquema del árbol de categorías y nombres de taxones que puede representarse como una entidad con relaciones unarias. 54
Figura 23. Esquema de la entidad NOMBRE y sus relaciones unarias. 56
Figura 24. Esquema de las entidades NOMBRE, AUTORTAXON Y CATEGORIATAXONOMICA con sus relaciones. 56
Figura 25. Comparación en el porcentaje de error al validar bases de datos con diccionarios y sin ellos. 58
Figura 26. Ventana principal de acceso al Sistema de Información BIOTICA. 61
Figura 27. Menú de la sección nomenclatural. 61
Figura 28. Pantalla del Catálogo de Categorías Taxonómicas. 62
Figura 29. Pantalla del Catálogo de Autoridades. 63
Figura 30. Pantalla Baja del Taxón Señalado. 64
Figura 31. Pantalla para mover un taxón de un nivel a otro. 64
Figura 32. Pantalla del Catálogo de Categorías Taxonómicas para ingresar nuevas categorías. 65
Figura 33. Pantalla de Alta Nuevo Taxón. 66
Figura 34. Pantalla de Captura y Modificación del Taxón para 66

ingresar nuevos nombres de taxones.
Figura 35. Pantalla de captura de PLATYPUS Versión 1.1. 68
Figura 36. Ejemplo de etiqueta mastozoológica. 78
Figura 37. Etiqueta para insectos preparados en alfiler. 78
Figura 38. Etiqueta usada para helmintos parásitos en frasco. 78
Figura 39. Etiqueta usada para helmintos en preparaciones microscópicas. 79
Figura 40. Etiqueta usada para peces. 79
Figura 41. Etiqueta usada para aves. 80
Figura 42. Etiqueta de herbario. 80

ÍNDICE DE CUADROS

Cuadro 1. Número de órdenes y familias de angiospermas en tres clasificaciones recientes (Tomado de Minelli, 1993). 4

Cuadro 2. Comparación entre dos programas: Access 2.0 de *Microsoft* y *Paradox® 5.0 para Windows*. 13

Cuadro 3. Ejemplo de los datos de plantas vasculares contenidos en la entidad Nombre. 39

Cuadro 4. Ejemplo de los datos de vertebrados contenidos en la entidad Nombre. 39

Cuadro 5. Ejemplo de los datos contenidos en la entidad AutorTaxon. 40

Cuadro 6. Categorías taxonómicas en la base de datos. Las categorías designadas por enteros en el campo nivel comunes a todos los taxones. 46

Cuadro 7. Categorías taxonómicas obligatorias y no obligatorias consideradas en la base de datos. 47

Cuadro 8. Sistemas de clasificación utilizados para cada taxón. 48

Cuadro 9. Número de especies/subespecies por taxón y sistema de clasificación. 48

Cuadro 10. Número de familias, géneros y especies totales para cada taxón. 49

Cuadro 11. Datos generales de las instituciones registradas en CONABIO que albergan colecciones científicas. 50

Cuadro 12. Tabla única para una clasificación de reptiles. 53

Cuadro 13. Comparación entre los modelos BIOTA y BIOTICA. 68

RESUMEN

Actualmente, las colecciones biológicas científicas se han enfrentado a la problemática del manejo de la gran cantidad de datos que albergan, ya que la información que poseen los especímenes de colección en las etiquetas que los identifican es muy variada. Sin embargo, todos los ejemplares de una colección deben poseer como rasgo común un nombre válidamente publicado y aceptado que pertenece a un sistema de clasificación particular. Para realizar más eficientemente las tareas de consulta y administración de las colecciones y su inmensa cantidad de información, se han realizado muchos intentos a nivel mundial para computarizar el acervo de las mismas a través de bases de datos. Esta tarea incluye el desarrollo de estándares para el ingreso de datos, con la finalidad de obtener una base de datos más confiable.

Los sistemas de clasificación no han estado al margen de la automatización de las colecciones. Los diccionarios o archivos de autoridades son listas de nombres válidos de taxones en formato electrónico que contienen la información jerarquizada de un sistema de clasificación para un determinado grupo de organismos. El diccionario consiste de los nombres correctos de los taxones de todas las categorías taxonómicas por lo que se le considera un estándar de control de calidad.

En el presente trabajo se desarrolló un modelo de datos para bases de datos relacionales con el objetivo de elaborar diccionarios de nombres de taxones que incluyen relaciones jerárquicas de acuerdo con un sistema de clasificación eje para plantas vasculares y vertebrados terrestres de México. El modelado se llevó a cabo siguiendo la metodología del modelo entidad-relación y se emplearon siete sistemas de clasificación con más de 35,000 nombres de taxones y 34 categorías taxonómicas diferentes.

Los diccionarios se acoplaron al modelo de datos BIOTICA de CONABIO, diseñado para el manejo de información de colecciones biológicas. El modelo diseñado permite que sea implementado independientemente del manejador de bases de datos y de la plataforma en particular, aunque el desarrollo del modelo se realizó en el manejador de bases de datos Access 2.0 de Microsoft.

Palabras clave: *diccionarios de autoridades, colecciones biológicas, bases de datos, nomenclatura, sistemas de clasificación.*

I. INTRODUCCIÓN

1.1 Colecciones biológicas científicas

Las colecciones científicas son el respaldo, la memoria y, de cierto modo, la culminación del esfuerzo que ha hecho el hombre por conocer y organizar su conocimiento de los organismos que lo acompañan poblando la tierra. Los museos y herbarios son centros de información no sólo porque los ejemplares biológicos que poseen son representativos de poblaciones y especies, y de esa manera aportan datos sobre la naturaleza y las características del ambiente en el que viven, sino también porque se acompañan de catálogos y bibliografía de diversos tipos sobre los ejemplares que ahí son atesorados (Peláez, 1994).

Las colecciones biológicas científicas representan un acervo especial de referencia para la descripción y análisis del mundo vivo. Los lugares en los que se encuentran colecciones científicas como los museos, son básicamente centros de información, y dadas las magnitudes de la información que incluyen, esta no puede ser manejada manualmente, por lo que la informática tiene mucho que hacer en estos centros. Además, las bases de datos biológicas, es decir, los conjuntos de datos sobre organismos o fenómenos biológicos que son almacenados en algún medio accesible por computadoras, son en sí mismas colecciones científicas (Crisci, *et al.* 1992).

Existen dos grandes tipos de colecciones, las de especímenes fijados y las de organismos vivos. El manejo de ambos tipos es completamente diferente. Los datos que se obtienen cuando un ejemplar es recolectado y preparado para formar parte de una colección permanente, corresponden a datos *in situ*, es decir, aquellos que se registran en el instante mismo de la recolecta. En cambio, un organismo vivo que se incorpora a una colección (por ejemplo, jardín botánico o zoológico) requiere de la recopilación de otra clase de datos, como son datos alimenticios, reproductivos, de comportamiento, ecológicos, etc., los cuales deben ser monitoreados durante toda la vida del individuo.

A pesar de que la mayoría de los datos entre estas colecciones difieren en gran medida, existen atributos que son comunes a ambas, tales como el nombre científico, que tiene la propiedad de ser universal, estable y único. En adelante, nos referiremos únicamente a las colecciones de organismos fijos.

Durante la catalogación de un organismo en una colección, se registra la información básica de cada ejemplar que se incorpora de manera permanente en la colección dentro de los catálogos cronológico o numérico, sistemático y geográfico (Ramírez-Pulido, *et al.* 1986).

La información básica del catálogo cronológico varía dependiendo del taxón del que se trate, pero todos y cada uno de los ejemplares debe poseer información indispensable como número de catálogo, familia, género, especie y subespecie, entidad federativa, localidad precisa, nombre del recolector, fecha de recolecta, altitud, número de recolecta y hábitat. La mayoría de estos datos son repetitivos en muchos casos y el capturarlos una y otra vez generan errores. Es posible entonces crear diccionarios controlados que beneficien tanto el acceso de los datos como la validación de la información.

Cada ejemplar que es alojado dentro de una colección lleva una o un juego de etiquetas que lo identifican. La relación que existe entre los ejemplares y sus etiquetas es inseparable y, de hecho, el extravío de cualesquiera de ellos se traduce en una pérdida del valor científico del especimen (Hawsks y Williams, 1986).

Existen datos que se comparten en los ejemplares que forman parte de una colección, sin embargo, las características del grupo hace que algunos otros deban ser considerados particularmente, es decir, se van haciendo más específicos, por lo que los formatos de las etiquetas se deben adecuar al grupo del que se trate; incluso en grupos muy grandes las formas de registro son muy variadas. En el anexo 1 se pueden observar diferentes tipos de etiquetas, las cuales presentan como característica general el nombre científico del ejemplar.

1.2 Los sistemas de clasificación nomenclaturales

La sistemática es la práctica y teoría de la clasificación biológica. Clasificar objetos es parte de la naturaleza humana y tiene su origen en la prehistoria. Las primeras sociedades humanas dieron nombre y reconocieron plantas y animales por razones prácticas, una clasificación básica para la supervivencia consiste en distinguir lo comestible de lo no comestible. Estas particiones de objetos puestos juntos dentro de grupos conceptuales y su práctica diaria es usada por todos nosostros. En la sistemática moderna, intentamos descubrir la diversidad total de los organismos. Al entender los procesos que producen esa diversidad, y clasificar los organismos, tratamos de expresar sus relaciones filogenéticas (historia evolutiva). Un aspecto importante de la sistemática es la asignación y uso de nombres taxonómicos-nomenclaturales (Zug, 1993).

El término clasificación es usado tanto para el proceso mismo de clasificar, como el producto de ese proceso, es decir, el sistema de clasificación. La clasificación biológica es el proceso de agrupar organismos y el subsecuente agrupamiento de estos dentro de grupos más grandes. Dadas las variaciones en la naturaleza, se pueden reconocer unidades discretas en el mundo biológico. Para poder hacer referencia a esas unidades, estas deben ser nombradas, y debe establecerse un sistema de nomenclatura aceptado. El sistema de clasificación contiene categorías (especies, géneros, familias, órdenes, clases, etc.) dentro de las cuales son colocadas las unidades, los grupos de unidades, y los grupos de grupos. El sistema de clasificación entonces necesita una jerarquía de categorías dentro de las cuales son colocadas las unidades y los grupos nombrados (Raddford, *et. al.* 1974).

Los sistemas modernos de clasificación intentan servir a dos necesidades: (1) presentar nuestro estado actual del conocimiento acerca de las relaciones (totalidad de similitudes y diferencias, sometidas a interpretaciones filogenéticas) de los organismos y (2) proveer en combinación con un sistema de nomenclatura, un sistema capaz de recuperar información acerca de los organismos (Raddford, *et. al.* 1974).

1.2.1 Sistemas de clasificación para plantas vasculares

Desde el siglo pasado, se han llevado a cabo los mayores intentos de clasificación de las plantas con flores, intentando integrar las similitudes y diferencias entre las plantas y su contexto evolutivo, y algunos autores han considerado implícitamente el contexto filogenético (Cronquist, 1993). De tal modo, las clasificaciones más actuales están diseñadas de manera que cada rama del árbol posea un orden identificado por un número (número filogenético). Por ejemplo: en el sistema de Takhtajan (1980), el número 1 corresponde a la familia Degeneriaceae, y para Engler (1964) esta familia tiene asignado el 40; así, el número 1 de Engler (1964) es la familia Casuarinaceae y el 40 de Takhtajan (1980) es Nymphaceae.

En el siguiente cuadro se muestra una comparación de los sistemas más recientes para angiospermas. Sin embargo, el cuadro no es útil para entender similitudes y diferencias mayores entre las clasificaciones por razones de sus continuos cambios, algunos de los cuales no siempre están claramente justificados (Minelli, 1993).

Autor de la Clasificación	Órdenes			Familias		
	Dicot	Monocot	Total	Dicot	Monocot	Total
Cronquist (1988)	64	19	83	322	66	388
Takhtajan (1987)	128	38	166	429	104	533
Dahlgren (1989)	86	25	111	371	104	475

Cuadro 1. Número de órdenes y familias de angiospermas en tres clasificaciones recientes (Tomado de Minelli, 1993).

La clasificación de Cronquist difiere menos de los proyectos tradicionales que las otras. Cronquist rechaza una aproximación cladística de la clasificación y explícitamente defiende el uso de taxones parafiléticos. Dahlgren hace manifiesto el punto de vista cladista como el criterio principal para desarrollar su sistema. Algunos años antes, cuando había un menor avance en el desarrollo de las mejores clasificaciones de las plantas con flores, parecían tener varias similitudes importantes entre los proyectos de Cronquist y de Takhtajan, por un lado; y las de Thorne y Dahlgren, por el otro. Los subsecuentes cambios hechos a Thorne y Takhtajan los hace más similares (Minelli, 1993).

En 1976, cuando fue presentada la segunda versión de este sistema, Thorne remarcaba que muchos taxones de los niveles más altos, que eran tradicionalmente aceptados dentro de plantas con flores y aún 'extensamente reconocidos, aunque algunas veces bajo el disfraz de otros nombres', son ciertamente polifiléticos. Ejemplos de ellos se observan en Amentiferae, Monochlamydeae, Dialypetalae y Sympetalae. Este hecho parece ser suficientemente reconocido y rechazado por todos los autores recientes, aunque sus clasificaciones aún incluyan muchos taxones parafiléticos (Minelli, 1993).

Las diferencias de opinión expresadas por Cronquist, Takhtajan y Thorne en sus respectivas clasificaciones no son tan grandes como parecen ser. Rara vez un autor circunscribe una familia al incluir taxones que otros asignan a lugares separados. Más frecuentemente, un grupo de familias puede ser asociado a un autor, pero quedaría disperso con respecto al concepto de otro autor. En muchas de esas instancias, las familias que se incluyen son extremadamente especializadas. Cuando se presentan diferencias de opinión, no hay una forma evidente de resolverlo. Así, Cronquist ha aceptado sólo 388 familias si se compara con las 533 reconocidas por Takhtajan, por tanto, la diferencia es significativa, aunque la gran mayoría de las familias aceptadas por Takhtajan están directamente asociadas con las familias adoptadas por Cronquist.

Las clasificaciones anteriores raramente o casi nunca incluían a los helechos y plantas afines, los cuales requerirían un tratamiento especial, obviamente diferente al de las angiospermas.

Para las Pteridofitas (helechos y plantas afines), se revisaron las clasificaciones de Tryon & Tryon (1982), Mickel & Beitel (1988) y Moran & Riba (1985). En el caso de las gimnospermas, algunas propuestas para la clasificación que se consideraron son las de Arnold (1948), Bold, *et. al.*(1980), Dogra (1980), Florin (1955), Flory (1936), Gaussen (1942-55), Hardin (1971), Hutchinson (1924) y Sporne (1974). Para las angiospermas, se revisaron: Engler (1964), Brummitt (1993) y Mabberley (1987).

1.2.2 Sistemas de clasificación para vertebrados

Sin duda alguna, uno de los principales problemas en la integración del conocimiento sobre los recursos naturales del país ha sido la falta de inventarios completos acerca de nuestra biota. En términos generales, los vertebrados de México están bien conocidos, en particular los terrestres. Sin embargo, los anfibios y los reptiles son los grupos menos conocidos, pues todavía están por describirse varias especies nuevas, y no existe una guía completa de los anfibios y reptiles de México. En el caso de los peces, el problema parece ser más acentuado, ya que se tiene muy poca información sobre este grupo en estudios ecológicos y de biodiversidad. Son pocos los trabajos sobre la diversidad biológica del país que mencionan la presencia de los mamíferos marinos en aguas patrimoniales mexicanas, en gran medida porque el estudio de la fauna mastozoológica marina en México realizada por investigadores nacionales es muy reciente y se remonta a los años cincuentas. Aunado a esto, se conoce que una porción muy alta de especies de las cuatro clases de vertebrados terrestres que habitan México son endémicas en el país: 60.7% de los anfibios, 53.7% de los reptiles, 7.6% de las aves, y 30.2% de los mamíferos. En conjunto, poco más de la mitad de las especies de anfibios y reptiles mexicanos es endémica al país (55.7%). Esto hace a la herpetofauna de México una de las más interesantes del mundo (Flores-Villela, 1993; Espinosa *et. al*, 1993; Salinas y Ladrón de Guevara,1993).

1.2.2.1 Sistemas de clasificación para mamíferos

A nivel mundial, la publicación de Corbet, G. & J. Hill (1991) es de las más reconocidas sobre mamíferos. La clasificación está basada en el trabajo de Simpson, publicado en 1945, bajo ciertas modificaciones en algunos taxones de Chiroptera y Muridae.

Antes de la publicación del trabajo de Ramírez-Pulido *et al.* en 1983, no existía un esfuerzo real por obtener una lista completa de las especies y subespecies de los mamíferos terrestres de México en la cual se incorporen cambios taxonómicos. Si bien es cierto que en trabajos posteriores se mencionaron los cambios taxonómicos y las nuevas categorías específicas y subespecíficas, la forma como se ha presentado esa información

no permite identificar con prontitud la diversidad mastozoológica mexicana. Posteriormente, en 1996, Ramírez-Pulido, *et al.* proporcionan una lista completa de las especies y subespecies de los mamíferos de México y recogen los cambios que se han publicado con posteridad.

Los cambios que se proponen incluyen: la secuencia como se presenta la información de los órdenes, separación de componentes de un mismo género, fusión de categorías genéricas, consideradas de prioridad tanto a nivel genérico como específico, revalidación de los nombres de especies y subespecies, incorporación de especies y subespecies en la sinonimia de otras y corrección de la nomenclatura de otras (Ramírez-Pulido, *et al.* 1996).

Ceballos, G. y H. Arita en 1996 recopilaron información de mamíferos del país a través del proyecto con clave A003 de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO): "Formación de una base de datos para el Atlas Mastozoológico de México", donde incluyen un sistema de clasificación el cual difiere en cierta medida con el de Ramírez-Pulido, *et al.* (1996).

En el trabajo de Salinas y Ladrón de Guevara (1993) aparece una lista sistemática de los mamíferos marinos mexicanos, en la cual se consideran tres órdenes y 41 especies.

1.2.2.2 Sistemas de clasificación para peces

Las bases del conocimiento de la ictiofauna continental mexicana fueron sentadas desde los años 40's por De Buen, y seguida por Álvarez y Miller, pero hasta 1993, aparece una lista sistemática con 506 especies mexicanas, recopilada por Espinosa *et. al.* (1993), quienes siguieron la clasificación propuesta por Nelson (1984). Esta cantidad de especies representa aproximadamente el 6.9% de los 8,411 peces dulceacuícolas del mundo (Nelson, 1984), cifra significativa en cuanto a diversidad, si se considera que el país cuenta con sólo 1.9% de las tierras emergidas en el planeta.

1.2.2.3 Sistemas de clasificación para reptiles y anfibios

En los últimos 16 años se han descrito o registrado por primera vez para el país aproximadamente 65 especies de anfibios y reptiles de México; se han hecho cerca de 130 cambios taxonómicos y nomenclaturales a nivel de especie y más de 50 por arriba del nivel de especie; se han publicado incontables cambios de distribución y nuevos registros de localidades (Flores-Villela, 1993).

Flores-Villela (1993) presenta una lista de especies de anfibios y reptiles válidas para México, elaborada con base en la revisión de literatura taxonómica reciente y los cambios que se han producido de acuerdo con los especialistas. Para la nomenclatura a nivel de órdenes y subórdenes se siguen los propuestos por Dundee (1989). En este caso se agregó el suborden Amphisbaenia, dado que este último autor no lo incluye. No se sigue la clasificación de órdenes ni de familias propuesta por Bour y Dubois (1984).

Para la mayoría de las especies de anfibios, tortugas, y cocodrilos, así como de los taxones por encima del nivel de especie, se siguen los criterios nomenclaturales propuestos por Frost (1985) para los anfibios y King y Burke (1989) para las tortugas y cocodrilos. Dubois (1985) da una lista de las categorías por arriba de género para los anfibios pero no se tiene información sobre algo similar para los reptiles en conjunto.

1.2.2.4 Sistemas de clasificación para aves

La clasificación adoptada para el diccionario de aves es la de la American Ornithologist's Union (AOU) de 1983, pero también se revisaron las de Jobling (1991) y Miller (1957). La clasificación de la AOU no considera la existencia de subespecies y con respecto a este criterio, México cuenta con más de 1,000 especies, pero para este trabajo, se incluyeron algunas de Norteamérica, siendo en total 2,099 correspondientes a 756 géneros y 93 familias.

1.3 La computarización de colecciones biológicas

Es un hecho que la computación se ha convertido en un recurso clave para el desarrollo de la ciencia moderna. Casi desde el mismo momento de su aparición, las computadoras fueron utilizadas en las ciencias naturales con resultados exitosos. Actualmente se emplean en una amplia gama de actividades, tales como cálculo, enseñanza, banco de datos, edición de textos, simulación de procesos naturales y control de experimentos (Crisci, *et al.* 1992).

La sistemática no ha estado al margen del fenómeno producido por la computación; por el contrario, el uso de computadoras obró como un verdadero catalizador en la transformación teórica y práctica de dicha disciplina. La aparición de una herramienta de cálculo tan poderosa y veloz trajo aparejada una verdadera revolución en los principios y propósitos de la clasificación de los seres vivos. Frente a esta situación ha surgido una gran cantidad y variedad de programas de computación (*software*) accesibles a la mayoría de los investigadores (Crisci, *et al.* 1992).

Actualmente el uso de las computadoras es aplicado a todos los campos, y las colecciones biológicas científicas no están exentas de adaptarse a las nuevas tecnologías (Martínez, H. *et. al.*, 1994 com. pers.).

En promedio, más del 70% del material depositado en las colecciones mexicanas está incorporado a colecciones de consulta, pero su accesibilidad es aún muy reducida. Los propios especialistas se ven a veces abrumados por la magnitud de la información de su colección al no contar con herramientas como las bases de datos, que les permitan un manejo ágil y eficiente de la información (Llorente y Koleff, 1997). Sin embargo, aquellas colecciones que ya poseen bases de datos se enfrentan a otros problemas que surgen cuando se realizan consultas que se deben frecuentemente a errores de captura al ingresar gran cantidad de datos; así, es de gran ayuda contar con diccionarios de nomenclatura del sistema de clasificación usado en la catalogación de los ejemplares.

Un sistema de información es el medio por el cual los datos fluyen de un usuario a otros, además proporciona servicios a todos los demás sistemas de una organización y enlaza todos sus componentes en forma tal que estos trabajen con eficiencia para alcanzar el mismo objetivo. Desde hace cerca de 20 años, se ha intentado elaborar un sistema de información sobre colecciones biológicas. Tales intentos se pueden agrupan

en dos grandes tendencias. Por un lado, están aquellos sistemas que tienen como objetivo principal la administración de las colecciones. Por otra parte, hay otros que buscan fundamentalmente la consulta remota y el intercambio de información entre las colecciones y su uso con fines de análisis.

1.3.1 Bases de datos relacionales

Un dato es la representación simbólica de cualquier fenómeno que el ser humano puede conceptualizar, es decir, se refiere a hechos concernientes a personas, objetos, eventos, u otros. Una colección biológica reúne grandes cantidades de datos, tanto de fuentes internas como externas, en cantidades que resultan agobiantes si se mantienen en su estado original, sin conceptualizar (Koleff, 1997).

La información es un conjunto de datos analizados que nos permite tomar decisiones. Estos datos, que han sido procesados y presentados de una forma adecuada para su interpretación, frecuentemente tienen el propósito de revelar tendencias o patrones (Koleff, 1997).

Una base de datos es un conjunto de información relacionada con un tema o propósito particular, la cual está bien estructurada y precisa y puede ser recuperada posteriormente (Microsoft Corporation, 1994a).

Una base de datos es un depósito de datos integrados, con redundancia controlada y con una estructura que refleja las interrelaciones y restricciones existentes en el mundo real; los datos han de ser compartidos por diferentes usuarios y aplicaciones, deben mantenerse independientes de estas y su definición y descripción han de estar almacenadas junto con los mismos. Los procedimientos de actualización habrán de ser capaces de conservar la integridad, la seguridad y la confidencialidad del conjunto de datos.

El diseño de una buena base de datos debe involucrar una buena estructura de las tablas que la conforman. La normalización es el proceso usado para asignar atributos a las entidades. La normalización reduce la redundancia en los datos y, por extensión, ayuda a eliminar datos anomalos que resultan de esa redundancia. La normalización no es eliminar la redundancia de los datos; en cambio, produce la redundancia controlada que permite ligar las tablas de la base de datos. El trabajo de normalización se lleva a

cabo a través de series de estados llamadas formas normales las cuales se traducen en relaciones de dependencia entre los atributos (Rob and Coronel, 1995).

1.3.2. Software para bases de datos

El *software* se define como "programas de computadora, procedimientos, y documentación y datos pertinentes a la operación de un sistema de cómputo" (Cota, 1994).

Desde los inicios de la utilización comercial de las computadoras en la década de 1950, las mejoras espectaculares en las tecnología de la producción de *hardware* han conducido a la fabricación de máquinas cada vez más poderosas, disponibles a precios cada vez más bajos. En cambio, la demanda social de *software* cada día más grande y compleja ha sufrido una atención deficitaria y los costos para obtener este *software* han crecido aún más rápidamente. Diversos autores han coincidido en señalar que, en promedio, las instalaciones gastan el 70% del presupuesto de las empresas para el desarrollo en el mantenimiento del software existente (Cota, 1994).

Para realizar una base de datos confiable, es necesario que el *software* y soporte técnico que se requiera sea igualmente confiable. Por tal motivo, la industria del *software* se ha dado a la tarea de crear innovaciones que beneficien a los usuarios y a los programadores. Sin considerar los requerimientos específicos, los productos de *software* que cumplan con las siguientes metas de calidad son los más fáciles de desarrollar y de adaptarse a necesidades no anticipadas (Cota, 1994):

- Adecuación. Un producto de *software* es adecuado en la medida que realice las funciones para las que fue construido de modo apropiado, consistente y predecible. En su sentido más amplio incluye la sensibilidad y tolerancia a fallas. Estas cualidades pueden lograrse más prácticamente si se tiene cuidado en el diseño y la construcción, más que en las pruebas y reparaciones subsecuentes.
- Utilidad. Un producto de *software* es útil en la medida que su estructura básica, su construcción y su documentación son útiles para el usuario al que está destinado.

- Costo-Eficiencia. Finalmente, un producto de *software* es costo-eficiente en la medida que sus costos de desarrollo y operación satisfagan las necesidades de los usuarios. A medida que el *hardware* disminuye su costo, la eficiencia de los sistemas tiende a ser menos crítica, de modo que el costo-eficiencia se expande para incluir tanto la reparabilidad como la modificabilidad. Entonces el costo-eficiencia global requiere que el producto sea fácilmente modificado y ampliado.
- Diseño. Esta actividad se realiza en dos partes: diseño general y diseño detallado. El diseño general es el establecimiento de la arquitectura global del sistema mediante la especificación de sus componentes organizados por niveles de abstracción, así como de las interfaces entre cada uno de estos componentes y niveles. Le sigue el diseño detallado de cada unidad de *software*, como una base para su implementación.

Una base de datos realizada en el *software* desarrollado por Microsoft Access, está compuesta por varios *objetos:* tablas, consultas (*queries*), formularios, informes o reportes, macros y módulos (Microsoft Corporation, 1994b).

Realizar una base de datos en Microsoft Access ofrece las siguientes ventajas:

- Es posible usar el nuevo comando de exportar en el menú de Archivo para crear rápidamente un archivo en formato de texto MS-DOS (.txt), formato de Microsoft Excel (.xls), o formato de texto de alta definición (.rtf), conteniendo la exportación de una tabla, forma, reporte o módulo.
- Exportar los resultados de una consulta a todos los formatos de archivos para tablas.
- Cuando se copian y pegan datos con una estructura particular desde una forma o tabla a un documento de Word para Windows o Microsoft Excel, Microsoft Access pega los datos con su formato (tales como fuente y color) intacto.
- Editar en el mismo lugar objetos de imagenes.
- Microsoft Access incrementa su compatibilidad con otras aplicaciones para Windows de Microsoft.
- Conexión a servidores de bases de datos *SQL* (*Structured Query Language*) (por ejemplo, Microsoft SQL Server o Oracle® Server) usando *drivers* de tipo *Open Database Connectivity* (*ODBC*). Después de que se instala el *driver ODBC* para las bases de datos *SQL*, se puede actualizar a la fuente de datos de cada base de datos *SQL* del tipo de acceso que se requiera. Entonces es posible pegar las tablas desde

la base de datos de la base de datos de Microsoft Access, o importar datos a y exportar datos desde la base de datos *SQL* (Microsoft Corporation, 1994[a]).

Además de las anteriores ventajas, Microsoft Access puede usar datos desde los siguientes formatos de base de datos y aplicaciones:

- Paradox 3.x y 4.x
- Microsoft FoxPro 2.0 y 2.5
- dBASE III y IV
- Btrieve (con archivos de datos definidos FILE.DDF y FIELD.DDF)
- SQL Base de datos, incluyendo Microsoft SQL Server, SYBASE SQL Server y Oracle Server
- Microsoft Access (otras bases de datos además de la base de datos abierta)

(Microsoft Corporation, 1994[a])

El siguiente cuadro presenta a manera de resumen una breve comparación entre Microsoft Access® 2.0 y Paradox® 5.0 para Windows™ para demostrar su utilización como una plataforma de fácil manejo para realizar este trabajo (www.microsoft.com).

Característica	Microsoft Access 2.0	Paradox 5.0 for Windows
Arquitectura	Microsoft Access 2.0 organiza automáticamente todos los objetos de la base de datos, permitiendo una navegación y transportación fácil entre ellos.	Paradox para Windows requiere de un archivo del sistema Microsoft-DOS para guardar los objetos organizadamente.
Transferencia básica de datos: importar y exportar	Microsoft Access 2.0 importa y exporta datos en todos los formatos populares. Los datos pueden ser importados a una tabla nueva o ya existente.	Paradox para Windows importa y exporta datos en muchos formatos de archivo populares, pero no en Lotus® 1-2-3® v. 3, Btrieve®, FoxPro, o Microsoft Excel 5.0.
Datos compartidos	Microsoft Access 2.0 lee y escribe directamente en dBASE, Paradox 3.X, 4.X; FoxPro 2.X, y Btrieve. Microsoft Access 2.0 se conecta con drivers ODBC para leer y escribir datos en Microsoft y SYBASE® SQL Servers. Además para otros servidores de datos populares (Oracle, Ingres, DB/2®, etc.)	Paradox para Windows lee y escribe directamente desde datos de Paradox. Los usuarios deben ordenar los *drivers* separadamente de SQL Link .
Manejo remoto de datos cliente /servidor	Todas las versiones de Microsoft Access incluyen la habilidad de leer directamente queries, y actualizar SQL back-ends. Además, se pueden crear queries a través de SQL que se ejecuten directamente desde el servidor. Los administradores pueden jugar con opciones como guardar contraseñas locales y otros. Las herramientas de actualización cliente/servidor de Microsoft Access 2.0's permiten desarrollar aplicaciones locales y "actualizarlas" desde el SQL Server, conservando datos, indexados, relaciones, integridad y reglas.	Paradox para Windows permite conectividad desde SQL pero debe obtenerse separadamente. Los administradores no pueden guardar contraseñas locales de Paradox/Windows y no tiene equivalente a las herramientas de actualización de Microsoft Access.

Cuadro 2. Comparación entre dos programas: Access 2.0 de Microsoft y Paradox® 5.0 for Windows.

El sistema de administración de bases de datos (DBMS) almacena y recupera información en una base de datos, y el sistema de administración de bases de datos relacionales (RDBMS) es aquel que almacena y recupera la información de acuerdo con relaciones que uno define. Se trata de una aplicación o programa informático (*software*), que posee Microsoft Access el cual permite organizar y analizar la información almacenada en tablas de una base de datos (Microsoft Corporation, 1994[b]).

Probablemente las bases de datos sean una de las aplicaciones más comunes de la informática actual. El gestor de bases de datos relacional es el programa que permite establecer relaciones complejas entre los datos de dos o más archivos al mismo tiempo. Las principales ventajas que encontramos son la rapidez de proceso y la posibilidad de realizar operaciones complejas o repetitivas disminuyendo los errores. Por esta característica es conveniente la utilización de gestores de bases de datos en aquellas aplicaciones en las que intervenga un considerable número de datos o en las que se tengan que hacer procesos repetitivos o muy complejos. Además no hay que contar únicamente con la utilización de la información. Uno de los trabajos más delicados y lentos a la hora de trabajar con las bases de datos se deriva de la introducción inicial de la información y de la actualización. Si se utiliza una base de datos con errores o no actualizada, es posible que nos encontremos con que las conclusiones obtenidas del proceso de información no sean todo lo verídicas y útiles que cabría esperar (Master Informática, 1992).

1.3.3 Modelos de bases de datos relacionales

Zellweger & Allkin (1993) presentan una a una las etapas de la metodología del modelado de bases de datos actual:

1. El modelo conceptual

Este nivel es el más aproximado a la percepción del mundo real. Intenta hacer una descripción formal y es esencialmente independiente de las limitantes de la implementación. Proporciona una excelente herramienta de comunicación y permite discutir acerca de los modelos de datos alternativos

2. El modelo lógico

Representa los datos por sí mismos y el tratamiento de esos datos dentro del contexto de una aplicación particular. Es una representación dinámica de los datos que refleja las necesidades y limitantes de la aplicación.

3. El modelo físico

Describe completamente el sistema de trabajo de la base de datos. Refleja las soluciones prácticas y constituye un plan de implementación detallado para un sistema de trabajo derivado del modelo lógico.

El modelo de entidad-relación (MER) es un modelo conceptual que ha sido ampliamente usado. Utiliza una representación del mundo real basado en dos clases de objetos: **entidades y relaciones**. Una técnica diagramática permite representar el modelo. "Entidades" son clases de objetos que podemos identificar y de sus miembros podemos guardar la información deseada. Diagramáticamente, las entidades son generalmente representadas por cajas. "Relaciones" son los nombres dados a asociaciones particulares existentes entre las entidades. Se representan diagramáticamente como diamantes, círculos, o simplemente como una línea vinculando las dos entidades asociadas con el nombre de la asociación escrita arriba (Zellweger & Allkin, 1993).

El modelo también permite la representación de información semántica como la naturaleza de cada relación de pares: esto es, uno-a-uno, uno-a-muchos o muchos-a-muchos. Esta información se puede representar usando una ramificación que indica un valor al final de la relación. Una relación muchos-a-muchos debe ser representada por una línea bidireccionada e igual al final (Zellweger & Allkin, 1993). Figura 1.

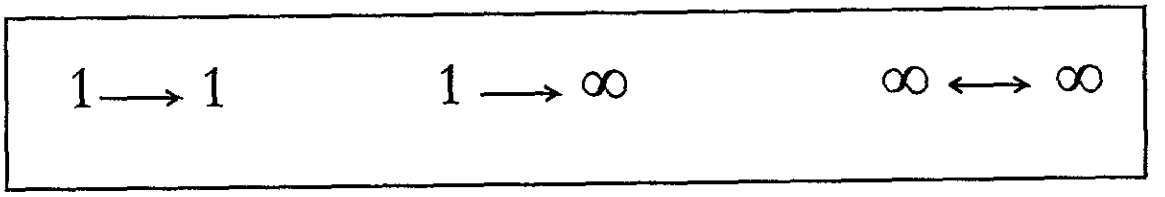


Figura 1. Tipos de relaciones entre entidades.

Los diagramas entidad-relación son de gran utilidad para explorar las relaciones existentes entre los taxones y sus nombres, y así iniciar la conceptualización del modelo.

Para el diseño de una base de datos se recomienda seguir los siguientes pasos (Microsoft Corporation, 1994a):

1. Determinar el propósito de la base de datos. Esto puede ayudar a decidir cuáles hechos se quieren almacenar en la base de datos.
2. Determinar las tablas que se necesitan. Una vez que se tiene claro el propósito de la base de datos, se puede dividir la información en sub-objetos separados. Cada sub-objeto será una tabla en la base de datos.
3. Determinar los campos que se requieren. Decidir que información se quiere guardar en cada tabla. Cada categoría de información en una tabla se llama un Campo y está desplegada como una columna en una tabla.
4. Determinar las relaciones. Observar a cada tabla y decidir cómo los datos en una tabla están relacionados con los datos de otras tablas. Añadir campos a tablas o crear nuevas tablas para clarificar las relaciones, tanto como sea necesario.
5. Refinar el diseño. Analizar los errores del diseño. Crear las tablas y adicionar unos pocos registros o muestras de los datos. Observar si se pueden traer los resultados esperados desde las tablas. Hacer reajustes al diseño como sea necesario
6. Diseñar consultas (*queries*) automatizadas.

1.3.4 Implicaciones y limitantes en el manejo y uso de bases de datos curatoriales

Los intentos de documentar las distribuciones de flora y fauna mundiales han sido frustrados por problemas prácticos. Entre esas dificultades están la compilación de bases de datos y la pregunta de cómo movilizar la gran acumulación de datos existentes. La compilación de bases de datos ambientales es lenta porque muchos proyectos adoptan una aproximación centralizada. Esto es, una agencia central adquiere los datos desde las personas que colectan y los ingresa a la base. La posesión de esta base de datos, la cual generalmente reside en una máquina simple, subordina al resto con los compiladores. El acceso a la base de datos está así severamente restringido. Esta aproximación involucra enormes problemas prácticos, los cuales incluyen altos costos de compilación,

duplicación de esfuerzos, derechos de autor, sensibilidad de los datos y obligaciones legales. (Green, 1994)

A continuación se presenta una variedad de limitantes las cuales pueden impedir el intercambio libre de datos. En muchos casos, tales limitaciones pueden ser perfectamente válidas (WCMC, 1997b):

- Inexistencia de lineamientos (ej. memorandums de entendimiento, políticas de intercambio e incorporación de datos).
- Derechos de autor (ej. *copyright*, confidencialidad).
- Procedimientos físicos onerosos (ej: recuperación,/compilación/edición/ copiado).
- Solicitudes insuficientes.
- Datos que no son de una calidad garantizada.
- Los usuarios tienen gran variedad de consultas, muchos de ellos con necesidades de información muy particulares (WCMC, 1997b)

Además, es imprescindible entender las responsabilidades de los usuarios con respecto al uso de la información, algunas de ellas son (WCMC, 1997b):

- Tomar decisiones racionales sobre las bases de la información de la red.
- Proporcionar advertencias sobre próximas necesidades de información, incluyendo presupuestos y catálogos.
- Proporcionar retroalimentación sobre la relevancia, sincronización y calidad de la información recibida
- Agradecer las fuentes en publicaciones y reportes

Muchos grupos humanos poseen conocimiento de naturaleza cultural, económica o científica el cual es de gran valor para la conservación y uso sustentable de los recursos biológicos. Esto ha sido construido por muchas generaciones y representa una inversión significativa en terminos de tiempo, dinero y esfuerzo intelectual. No sorprende, por lo tanto, que el flujo de información entre diferentes segmentos y niveles de la sociedad resulte frustrado por barreras políticas, organizacionales y personales (WCMC, 1997b).

Otra limitante en el manejo de los datos, es que estos deben ser protegidos del para el acceso por usuarios no autorizados. Consecuentemente se debe suministrar a los datos de las siguientes características (Rob and Coronel, 1995):

- *Seguridad física.* Permite el acceso fisico en áreas abiertas únicamente a personal autorizado.

- *Seguridad por contraseña.* Permite el acceso a usuarios autorizados y generalmente es reforzado con una clave de entrada.
- *Derechos de acceso.* El cual es establecido a través del usoi del software por ejemplo, algunas operaciones (creación, actualización, eliminación, etc.) son restringidas sobre objetos predeterminados como tablas, consultas y reportes.
- *Usuarios autorizados.* El DBMS generalmente verifica las violaciones de acceso a través de auditorias.
- *Datos ocultos.* Son usados para que los usuarios no autorizados no puedan romper algunas de las reglas de seguridad de la base de datos.

1.4 Bases de datos para consultar y administrar colecciones

Se han identificado principalmente dos tipos de bases de datos útiles para el trabajo de los taxónomos, las curatoriales –que almacenan información relacionada con un conjunto de especímenes depositados en una colección– y las taxonómicas (descriptivas) –que concentran información relacionada con especies–. Una base de datos de la mayor importancia para la taxonomía es la nomenclatural. Una de las características más organizadas de la taxonomía es el sistema de nomenclatura, las reglas de denominación de las especies están claramente establecidas, y son independientes de los diversos criterios de delimitación de las especies. Entre las bases de datos auxiliares de las taxonómicas y las curatoriales se encuentran en primer lugar las geográficas –que contienen información sobre localidades de recolecta–. A partir de las bases de datos curatoriales se pueden generar otras como las biogeográficas –que contendrían información acerca de la distribución de las especies con respecto a sistemas de clasificación biogeográfica establecidos y cartografiados (Peláez, 1994).

Existen al menos tres tipos principales de diseño para bases de datos que se han aplicado al desarrollo de sistemas de cómputo para colecciones biológicas: tablas únicas (*flat-file*), bases de datos relacionales y orientadas a objetos. Este trabajo considera la construcción de diccionarios para aplicarse particularmente a bases de datos relacionales.

Un modelo básico de las bases de datos relacionales para colecciones que consideramos en este trabajo es el siguiente (figura 2):

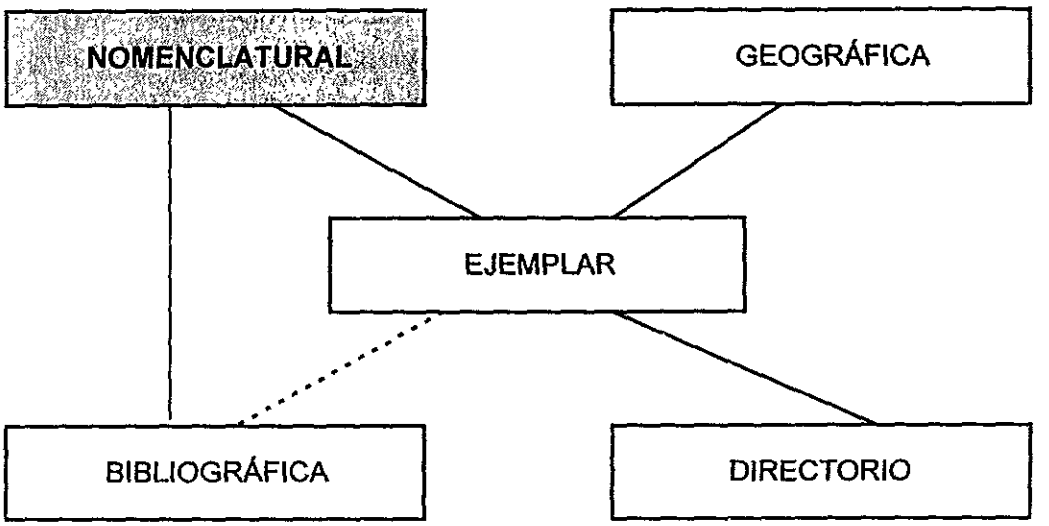


Figura 2. Entidades básicas de una base de datos relacional curatorial cuyo centro es la información del ejemplar. El objeto de estudio de este trabajo es la entidad nomenclatural.

Un área del manejo de bases de datos biológicas, que hasta ahora se ha comenzado a optimizar efectivamente es la organización de datos de especímenes e información de niveles taxonómicos dentro del contexto de la jerarquía taxonómica, permitiendo múltiples clasificaciones alternativas y la ordenación de datos nomenclaturales a todos los niveles jerárquicos. Las nuevas clasificaciones están generalmente construidas por conceptos taxonómicos compartidos y cambiantes de clasificaciones existentes. El desarrollo de un modelo de información aplicado a la taxonomía debe permitir la incorporación y ajuste de los conceptos de taxón de clasificaciones existentes cuando se proponen nuevas clasificaciones. Cualquier modelo de procesamiento de la información para clasificaciones debe ser capaz de capturar los conceptos taxonómicos (nombres y categorías) y las relaciones entre ellos mismos como de sus entidades de datos (Zhong *et al.*, 1996).

1.4.1 Esfuerzos para computarizar colecciones

La información biosistemática y su producción puede ser agrupada en 3 diferentes clases: físicas (el uso actual de colecciones curatoriales biológicas), impresas (revisiones taxonómicas, monografías, floras y faunas) y electrónicas (discos, CD-ROM, redes electrónicas) (Nielsen & West, 1994). Dentro de las impresas, se encuentran las listas de referencia de especies, las cuales contienen información mínima: generalmente el nombre binomial correcto, autor y año de publicación en un orden sistemático; en muchos casos incluyen información general sobre distribución geográfica y ecológica, aunque no permiten la identificación, la disponibilidad del nombre correcto del taxón permite accesar a la literatura, a bases de datos y a la comunicación, por esto, su frecuencia de uso es alta y la comunidad de usuarios es amplia (Nielsen & West, 1994).

La vasta cantidad de información contenida en las colecciones puede ser movilizada y hacerse más disponible. La tecnología de las computadoras permite el desarrollo de sistemas de información de alta calidad de todos los organismos, que incluya su taxonomía, nomenclatura y clasificación organizado en forma de bases de datos, con capacidad de generar *check-lists* (listas de referencia) y catálogos. Del mismo modo, la tecnología de bases de datos computarizadas hace posible la captura, examen y extracción de información de datos de la etiqueta de especímenes de colecciones biológicas. Las bases de datos descriptivas computarizadas, en este caso, las bases de datos taxonómicas, hacen el mismo proceso descriptivo más eficiente y conciso, y permiten la identificación en línea interactiva soportada por gráficos amigables al usuario. La tecnología de cómputo es capaz de almacenar y hacer uso efectivo de grandes bases de datos taxonómicas; tales bases de datos pueden también ser compartidas en redes electrónicas (Nielsen & West, 1994).

Se requiere mejorar la eficiencia en el manejo de las colecciones y evitar la pérdida de estándares curatoriales, así, es importante adoptar y desarrollar nuevos métodos y hacer uso de nuevos materiales. La información de las colecciones está disponible generalmente sólo con el acceso físico a la colección dado que sólo una muy pequeña proporción de los atributos de las colecciones han sido incorporadas en una base de datos computarizada (Nielsen & West, 1994).

Hacer más efectivo el progreso en el manejo de la colección, la investigación taxonómica y las bases de datos, incluye el uso de *check-lists* y catálogos, siendo las palabras clave un "protocolo estable" y un "control de calidad". Esos aspectos críticos garantizan que todos los recursos invertidos sean efectivos y que cualquier recurso adicional agregue valor a los datos. Con referencia a descripciones de nuevos taxones, un porcentaje significativo de nuevos nombres han sido después reconocidos como sinónimos. Puede ser interesante tener conocimiento de cuántos controles de calidad son implementados por las instituciones y *journals* que publican los artículos con un alto nivel de sinonimia (Nielsen & West, 1994).

Se tiene la posibilidad de comunicarse sin ambigüedad al hacer referencia a los organismos vivientes por el uso internacional de un nombre binomial, para el que existen reglas internacionales que regulan el uso y aplicación de dichos nombres. Con la tecnología de computadoras disponible actualmente no hay limitaciones para producir listas de referencia de especies y catálogos de todos los nombres introducidos en la literatura y su estatus respectivo (Nielsen & West, 1994).

Las bases de datos distribuidas de dominio público ofrecen soluciones para muchos de los problemas que obstaculizan la compilación de información sobre biodiversidad. Las herramientas de *software* necesarias para recuperar de la red el texto asociado e imágenes está totalmente disponible de forma gratuita en Internet. Las características esenciales de una base de datos distribuidas de dominio público son: acceso universal, mantenimiento descentralizado de los componentes de los datos, y el uso de formatos estándares de ingreso y actualización de datos. Los protocolos de validación deseables incluyen revisiones cruzadas para probar consistencia, uso de etiquetas de control de calidad y ligas de datos a los especímenes y, como consecuencia, genera una necesidad de nuevas convenciones de nomenclatura y campos clave registrados y otros índices taxonómicos secundarios (Green, 1994).

La idea de computarizar tanto diccionarios taxonómicos como información de colecciones científicas fue desarrollada formalmente desde la década de los 70's. A partir de entonces ha habido una diversidad de iniciativas de diferentes instituciones en el mundo.

Entre los esfuerzos mundiales que se han dado para eficientizar el manejo y la computarización de colecciones se encuentran los siguientes ejemplos, la mayoría de los cuales están disponibles a través de sus páginas de Internet:

COLECCIONES BOTÁNICAS

FINBIN - Finnish Biodiversity Information Network	http://www.metla.fi/biodiversity/
(CBIN) Canadian biodiversity information network	http://www.cbin.ec.gc.ca/
The Biodiversity and Biological Collections Web Server	http://www.leeming.wa.edu.au/WWWTour/local/science/bio.htm
TAMU Biology Herbarium via wwwRMG	http://csdl.tamu.edu/FLORA/tamuherb.htm
The Missouri Botanical Garden	http://www.mobot.org/
Flora of Australia (ERIN)	http://kaos.erin.gov.au/land/conservation.html
(SMASCH) Specimen Management System for California Herbaria	http://www.calacademy.org/smasch.html
Internet Directory for Botany "Checklist and floras,Taxonomical"	http://www.sun.ac.za/local/academic/natural/botany/idb/subject/botflor.html
Aquatic Plant Database	http://aquat1.ifas.ufl.edu/welcome.html
Carnivorous Plant Database	http://www.hpl.hp.com/botany/public_html/cp/html/actualcp.htm
TROPICOS Database - Missouri Botanical Garden. W3 TROPICOS	http://www.agnic.nal.usda.gov/agdb/w3tropic.html http://mobot.mobot.org/Pick/Search/pick.html
PLANTS National Database	http://trident.ftc.nrcs.usda.gov/
FLORIN Information System Home Page (Russia)	http://www.florin.ru/florin/

COLECCIONES HERPETOLÓGICAS

California Academy of Sciences Herpetology Catalogo vía gopher	http://www.calacademy.org/research/herpetology/
Northern Prairie Publications Database	http://www.npwrc.usgs.gov/
Crocodilian species list	
Texas Natural History Collections, The	http://www.utexas.edu/depts/tnhc/

Herpetology OnLine Database vía Gopher	
Swedish Museum of Natural History	http://www.nrm.se/about.html.en
Herpetology Collection Search Database	

COLECCIONES ICTIOLÓGICAS

The Biodiversity and Biological Collections	http://www.calacademy.org
Web Server Fish Searches	
Florida Museum of Natural History	http://www.flmnh.ufl.edu/
Ichthyology Collection Database	
University of Washington Fish Collection	http://www.ed/
Texas Natural History Collections fish	http://www.utexas.edu
database vía Gofer	

COLECCIONES MASTOZOOLÓGICAS

Yale Peabody Museum Mammalogy	http://www.peabody.yale.edu/collections/mam/
Collections vía Gopher	History.html
University of Alaska Museum's Mammal	http://www.alaska.edu/
Collection	

COLECCIONES ORNITOLÓGICAS

Cornell Ornithology Collection vía gopher	http://www.cornell.edu/Campus/CUFacts/ CUForni.html

1.4.2 El modelo de datos BIOTICA

A lo largo de casi 20 años se han replanteado los modelos sobre los que se construyeron algunas de las anteriores bases de datos y se ha comenzado a configurar un esquema de uso cada vez más estandarizado. BIOTICA, desarrollado en la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), es un modelo de datos para el manejo de la información curatorial, biosistemática, geográfica y bibliográfica que incluye rutinas de validación e interfaces para captura, reportes, visualización y análisis. BIOTICA es el esqueleto del SNIB (Sistema Nacional de Información sobre Biodiversidad), sistema en el que se integran datos sobre las especies con información complementaria de muy diversas fuentes (figura 3) (CONABIO, 1998).

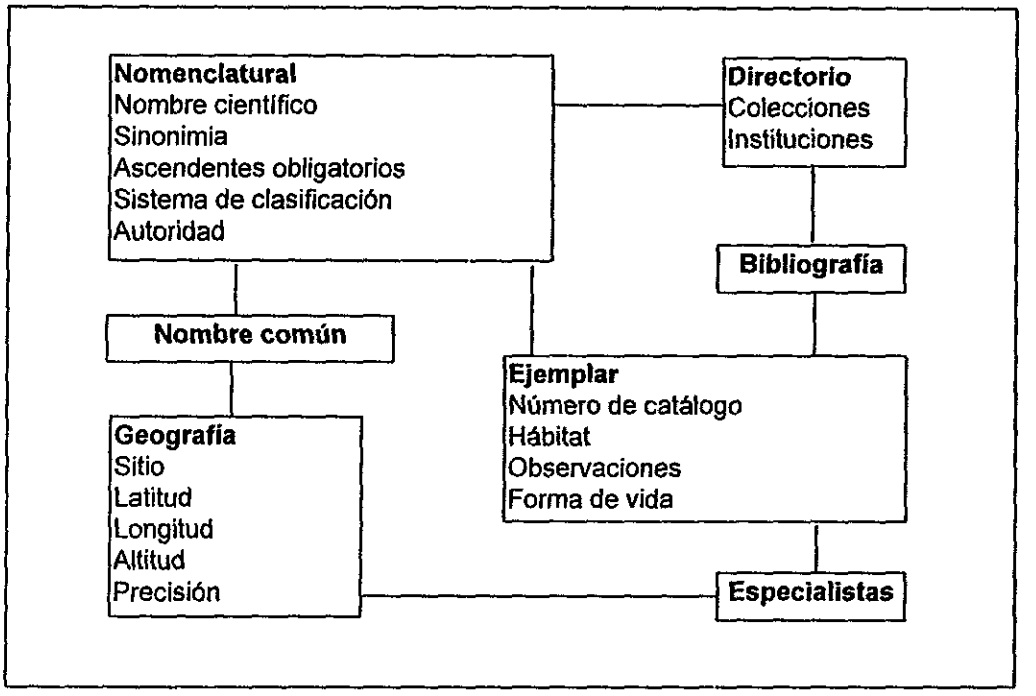


Figura 3. Tablas básicas del modelo de datos BIOTICA (CONABIO, 1998).

BIOTICA se ha desarrollado para ser un modelo de datos que pueda ser utilizado sobre diferentes manejadores y plataformas. Las versiones más comunes son dos, una sobre un manejador de bases de datos para usuarios en computadoras personales (ACCESS), y otra para integrar la información en el SNIB sobre un manejador más robusto (SYBASE), instalado en estación de trabajo. El modelo BIOTICA tiene la intención de automatizar búsquedas y consultas sobre colecciones de especímenes biológicos y es capaz de enlazarse entre esas diferentes versiones a través de interfaces.

1.4.3 Diccionarios de autoridades y catálogos controlados

Al ingresar nombres científicos a una base de datos, el usuario se enfrenta a un caso común de captura difícil debido a su escritura, y su captura recurrente ocasiona errores ortográficos, por ejemplo:

- Caesalpinaceae
- Caesalpiniaceae

lo mismo ocurre con los nombres de autores:

- Maxim. ex Franch. & Sav.
- Maxim. ex Francn. & Sav.

En los dos casos existe un error que la computadora detecta inmediatamente como series de caracteres diferentes, lo que ocasiona que se "duplique" la información: en lugar de tener una familia, la discrepancia en la escritura ocasiona que se cuantifiquen como dos datos diferentes. Para ambos ejemplos, es difícil decidir cuál es el nombre correcto, e incluso visualizar el error sin la consulta de índices reconocidos, bibliografía especializada, etc.

La solución sería contar con un diccionario precargado en la base de datos al cual se pueda recurrir cada vez que se captura un ejemplar nuevo, simplemente seleccionando la opción, lo cual evita teclear errores y aumenta la eficiencia en la captuta de datos al hacerse más rápido (un solo "*click*" del ratón o *mouse*). También es de gran ayuda cuando tenemos una base de datos elaborada y deseamos verificar la información al contrastar el diccionario con nuestra base de datos.

Los archivos o diccionarios de autoridad taxonómica, son listas de los nombres aceptados de un taxón, es decir, son archivos que contienen la información jerarquizada dentro de un sistema de clasificación para un determinado grupo de organismos. El diccionario consiste de los nombres correctos de los taxones de todas las categorías taxonómicas. En el caso de los nombres de las especies debe incluir además la información del autor y año.

En la perspectiva del SNIB es particularmente significativa la taxorreferencia, la decir, la expresión de una hipótesis sobre una posición actual del especimen dentro de un sistema taxonómico. La taxorreferencia está sujeta a errores e impresición, por lo tanto, esta información debe ser actualizada y modificada periodicamente por especialistas. Los nombres científicos son el lenguaje esencial para comunicarse acerca de la biodiversidad y forman parte de la taxorreferencia (Soberón and Koleff, 1998).

Una tarea importante es la necesidad de llevar a cabo un control de calidad de los datos. Sin él, no es posible proporcionar información confiable. La información de los especimenes en las bases de datos es obtenida de proyectos realizados por grupos de investigación muy diversos, así, el proceso de control de calidad es necesario porque los datos presentan numerosos problemas, tales como determinaciones mal realizadas, inestabilidad taxonómica, entre otros. Es posible detectar un gran número de esas inconsistencias si los registros son verificados contra archivos de autoridad (*authority files*), como es el caso de los catálogos taxonómicos para validar escritura correcta y sinonimia. El proceso de control de calidad (figura 4) proporciona bases de datos más pequeñas pero más confiables (Soberón and Koleff, 1998).

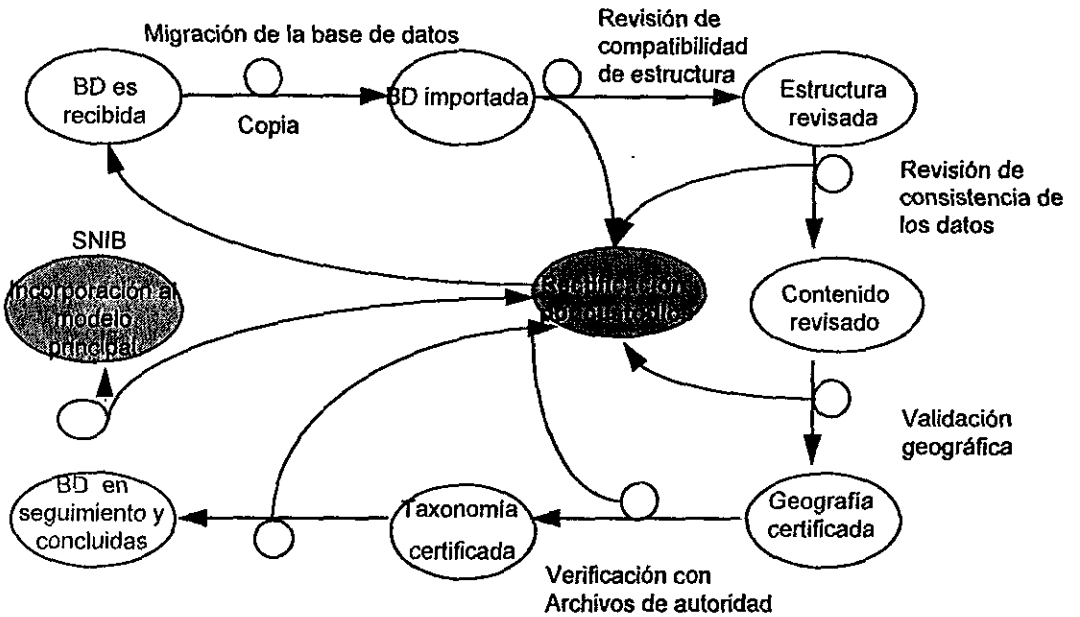


Figura 4. Proceso de control de calidad llevado a cabo en los datos que son incorporados al SNIB (Soberón and Koleff, 1998).

La conformación de un banco de archivos de diccionarios de autoridades y vocabularios controlados ha incluido, en un inicio, la elaboración de los catálogos para los grupos más representativos en las bases de datos del SNIB: plantas vasculares y vertebrados (figura 5).

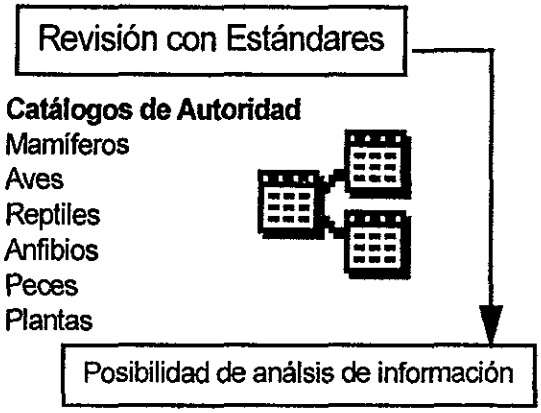


Figura 5. Catálogos desarrollados para llevar a cabo el control de calidad de las bases de datos del SNIB.

En el modelo de BIOTICA, el presente diccionario funge como el esqueleto de la tabla NOMBRE (figura 6).

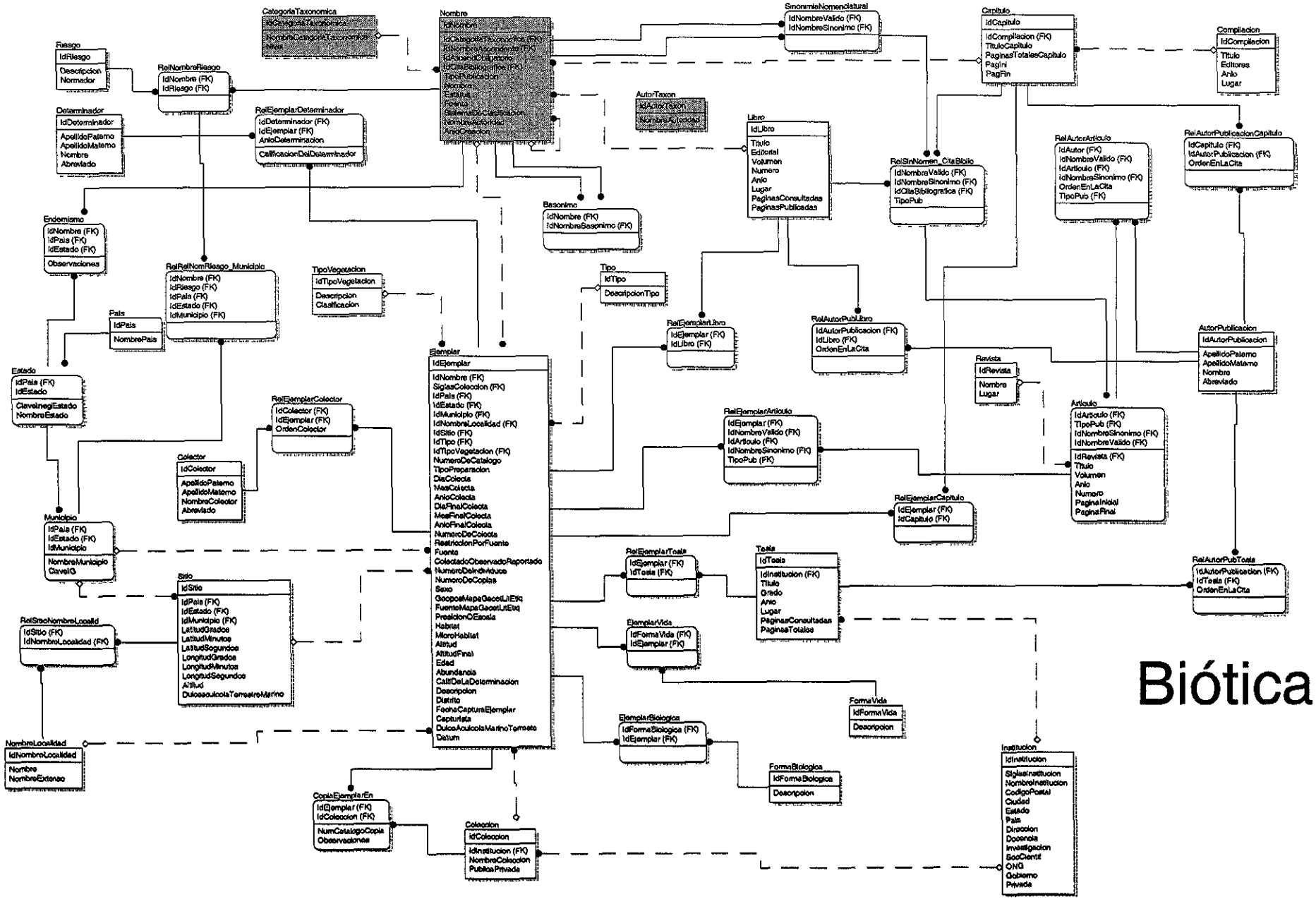


Figura 6. Modelo de datos BIOTICA

II. OBJETIVOS

1. Elaborar diccionarios de nombres de taxones, incluyendo relaciones jerárquicas de acuerdo con un sistema de clasificación eje para plantas vasculares y vertebrados terrestres.

2. Acoplar los diccionarios al modelo de datos BIOTICA de CONABIO, diseñado para el manejo de información de colecciones de herbarios, museos zoológicos y otras colecciones biológicas, independientemente del manejador de bases de datos y de la plataforma en particular.

III. MÉTODO

1. Se recopiló información sobre sistemas de clasificación para cada grupo, incluyendo la búsqueda de información de los sistemas por medio de la consulta a especialistas y en literatura.

2. Se obtuvieron diccionarios impresos sobre sistemas de clasificación de cada taxón.

Esto involucró la adquisición de publicaciones escritas y medios electrónicos de las clasificaciones que resultaron seleccionadas para cada taxón.

3. Se elaboraron tablas y formularios para la captura de nombres de taxones supraespecíficos y específicos y se depuraron ortográficamente.

Se llevó a cabo la captura de los nombres de los taxones, incluyendo autor y año de publicación.

4. Asignación de categorías.

Una vez que los nombres aceptados fueron capturados se procedió a asignarles la categoría taxonómica correspondiente. Para tal efecto, fue creada una tabla que contenía todas las posibles jerarquías y un identificador con el cual se podría llevar a cabo un vínculo entre ambos.

La entidad que contenía los nombres de taxones incluía en su estructura una relación unaria al hacer referencia a ella misma. Las relaciones unarias entre los diferentes niveles se realizaron con ayuda de formatos de pantalla, los cuales permitían visualizar la categoría superior de un nivel determinado y asignaban el identificador que realizaba el vínculo (ej.: para un género ***X***, era posible observar los nombres de todas las familias capturadas y elegir la familia ***Y*** a la que correspondiera).

En la figura 7 se presenta un esquema del árbol de nombres que contiene la tabla. Así, el nombre 6 tiene asignado una categoría III (clase) y su nombre ascendente es 3 con categoría II (división), pero al mismo tiempo 6 es ascendente de 8, 9 y 10 que tienen la categoría IV (orden).

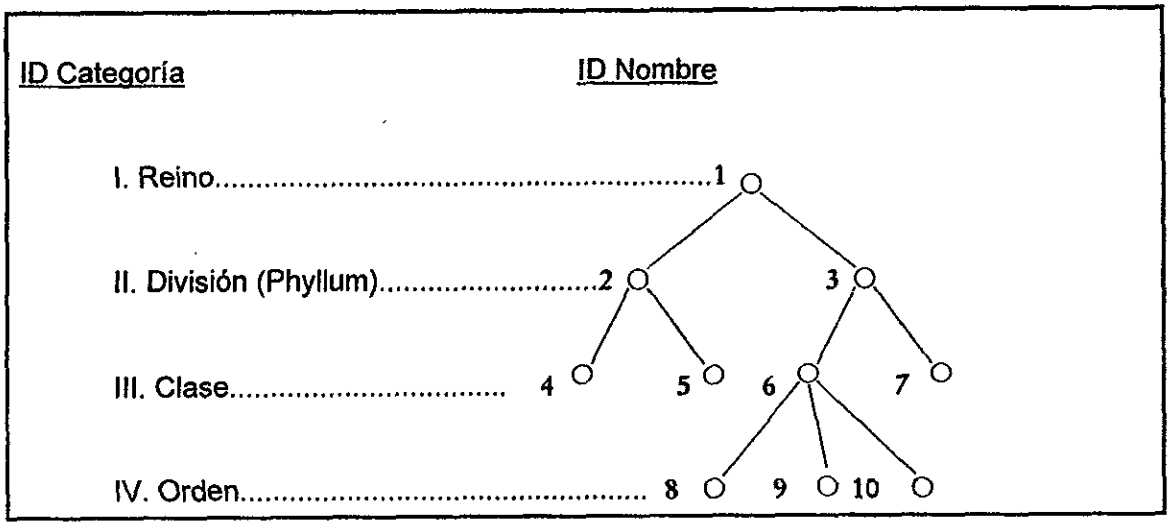


Figura 7. Esquema del árbol de categorías ascendentes obligatorias y nombres incluído en la entidad con relaciones unarias.

Supongamos que existiera una categoría "intermedia" entre clase y orden, por ejemplo, subclase. En este caso la llamaremos **III.*X***, de tal manera que 8, 9 y 10 tendrían como ascendente a 6.*X* y a 6 como ascendente obligatorio (figura 8).

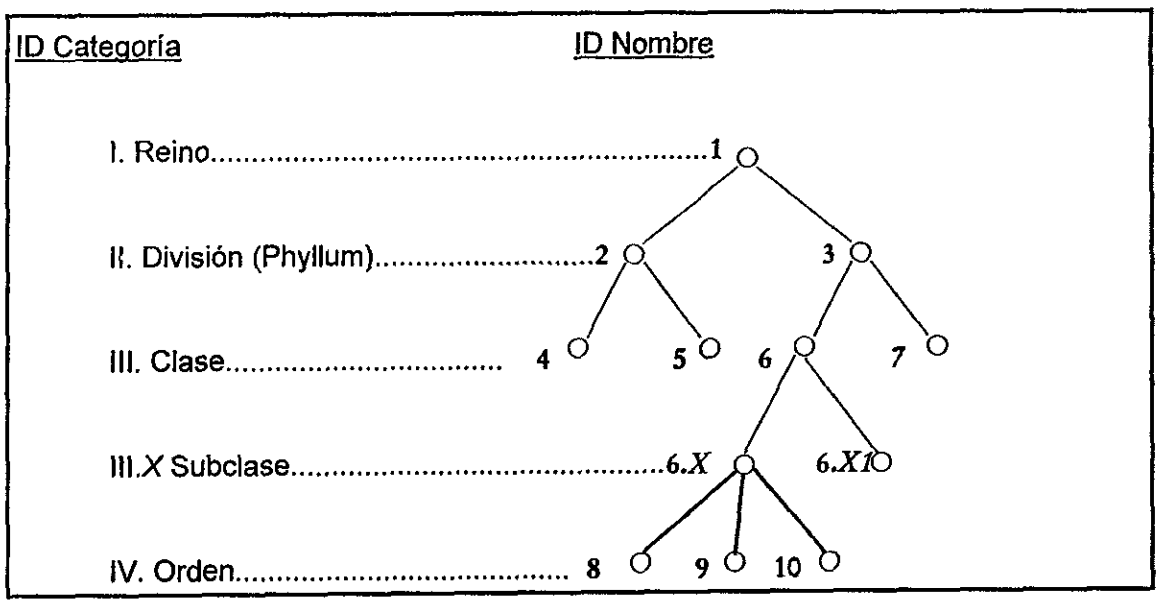


Figura 8. Esquema del árbol de categorías ascendentes y nombres incluído en la tabla recursiva.

Cada nivel del árbol se comporta de igual manera y cada combinación del identificador del nombre y el identificador del ascendente es única.

5. Depuración y prueba.

Las tablas fueron depuradas revisando consistencia, integridad, uniformidad y congruencia de la información normalizándolas por medio de consultas que permiten encontrar duplicados, errores de captura, etc.; además se sometieron a prueba a través de consultas con las cuales es posible recuperar toda la clasificación desde reino hasta especie.

6. Incorporación a BIOTICA.

La base de datos terminada se incorporó al modelo BIOTICA formando el esqueleto de la entidad NOMBRE, parte fundamental de la información curatorial de un especimen, incluyendo además las entidades CATEGORIATAXONOMICA y AUTORTAXON.

En la figura 9 se muestra un diagrama de flujo del método empleado en este trabajo.

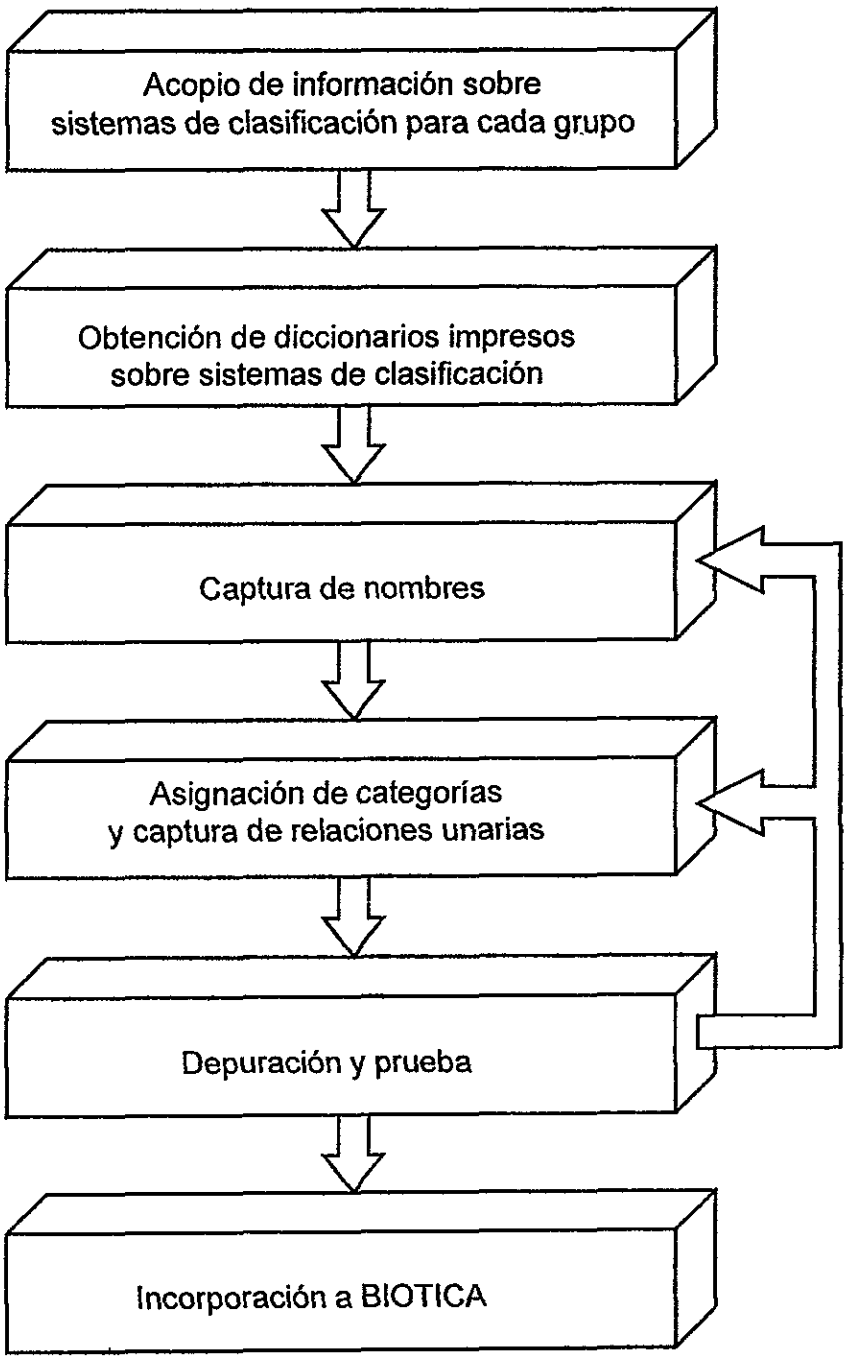


Figura 9. Representación esquemática del método.

IV. RESULTADOS

4.1 Sistemas de clasificación para plantas vasculares y vertebrados

Para las Pteridofitas (helechos y plantas afines), se revisaron las clasificaciones de Tryon & Tryon (1982), Mickel & Beitel (1988) y Moran & Riba (1985). Se consultó a especialistas para decidir cuál sistema de clasificación debería incluirse como diccionario de autoridades. El Dr. Ramón Riba y Nava Esparza, investigador de la Universidad Autónoma Metropolitana Unidad Iztapalapa, realizó los siguientes comentarios respecto a las tres clasificaciones:

a) Moran & Riba 1995 (Flora Mesoamericana, vol. 1, Pteridofitas) es el tratamiento florístico más amplio que se ha publicado hasta la fecha y es el resultado del acuerdo de cerca de treinta pteridólogos reconocidos en su campo, los que han aceptado esta clasificación. Por su misma estructura, será el instrumento de trabajo más usado por los pteridologos interesados en plantas tropicales. Esta fue la clasificación utilizada para la construcción del diccionario.

b) Tryon & Tryon 1982 (Ferns and Allied Plants) aunque de indiscutible calidad, no es un tratamiento florístico y no está al alcance de mucha gente interesada en el tema, además de que algunos grupos (p.e. helechos polipodiáceos) no van de acuerdo con las clasificaciones y puntos de vista más recientes.

c) Mickel & Beitel 1988 (Pteridophyte Flora of Oaxaca, México) presenta una clasificación basada en el sistema de Crabbe, Jermi & Mickel (A New generic sequence for the pteridophyte herbarium. Brit. Fern Gaz. 11: 141-162) que si bien incluye los géneros de todo el orbe básicamente descansa en los géneros del Viejo Mundo por lo que, para los taxa americanos, no concuerda con los conceptos aplicables a estos.

En el caso de las gimnospermas, algunas propuestas para la clasificación que se consideraron son las de Arnold (1948), Bold, *et. al.*(1980), Dogra (1980), Florin (1955), Flory (1936), Gaussen (1942-55), Hardin (1971), Hutchinson (1924) y Sporne (1974). De las anteriores se eligió la de Bold, *et. al.* (1980), después de haber consultado a diversos especialistas en coníferas del país .

Para las angiospermas, se revisaron: Engler (1964), Brummitt (1993), y Mabberley (1987).

Las dos últimas son las más utilizadas por los investigadores de los herbarios de México y el mundo. La escuela del Royal Botanical Garden Kew utiliza la clasificación de Brummitt (1993), y los principales herbarios de Estados Unidos están organizados conforme a la clasificación de Cronquist (Mabberley, 1987) o modificaciones de él. Se consultó la opinión de la entonces curadora del principal herbario del país (Herbario Nacional de México, MEXU), la Dra. Patricia Dávila, quien mencionó que ambos sistemas de clasificación son los más consultados y presentan diferencias significativas en el tratamiento a nivel de familia.

Las clasificaciones que se adoptaron para los mamíferos terrestres fueron la de Ramírez-Pulido, *et al.* (1996) como el sistema de clasificación eje, y como clasificación alterna se tomó la de Ceballos y Arita (1996) (utilizado por instituciones de la UNAM). En ambos casos se incluyó la propuesta para mamíferos marinos de Salinas y Ladrón de Guevara (1993).

A nivel internacional, se han producido varios archivos de autoridad para peces, tales como *FishBase* y *The Catalog of Fishes Project*. Cabe destacar que este último requirió 14 años de investigaciones y un millón de dólares para su construcción e implementación. A pesar de que está a disposición de cualquier usuario, el conocimiento de las especies mexicanas en él, es aún muy limitado. Se decidió entonces utilizar la clasificación de Espinosa *et. al.* (1993) para niveles de familia e inferiores, y para niveles superiores la de Nelson (1984).

El sistema de clasificación que se capturó para los taxones de reptiles y anfibios fue el de Flores-Villela (1993), que incluye 995 especies agrupadas en 199 géneros y 51 familias, debido a que es el único trabajo disponible que contempla una revisión específica para las especies mexicanas.

En el caso particular de las aves, dado que existe una organización internacional que regula la nomenclatura (AOU, American Ornithologit's Union) y los cambios nomenclaturales de las especies, se utilizó su última edición. A finales de 1996 se creó la *North American Biodiversity Information Network* (NABIN), donde actualmente el Dr. Adolfo Navarro de la UNAM, en coordinación con el Dr. A. Townsend Peterson del Museo de Historia Natural de la Universidad de Kansas, se encuentran actualizando el archivo de

autoridad con las modificaciones más recientes al sistema de clasificación y equivalencias a otros sistemas de clasificación (NABIN, 1997).

4.2 Descripción del modelo de la base de datos

Se elaboró una base de datos de tipo relacional en el manejador Access 2.0 de Microsoft. El modelo contiene una entidad con relaciones unarias con los nombres de los taxones de las clasificaciones eje y alternativas (llamada NOMBRE), una entidad de nombres de autores (AUTORTAXON) y una de categorías taxonomicas (CATEGORIATAXONOMICA). En la figura 10 se pueden observar las entidades y las relaciones que componen a la base de datos.

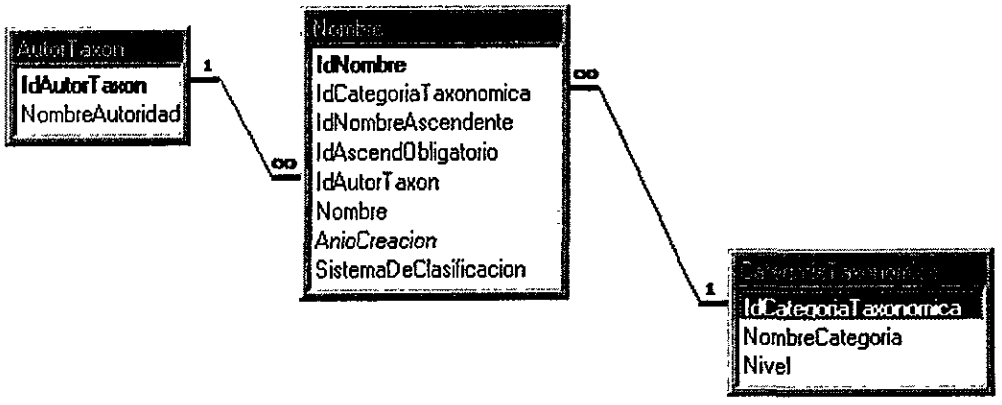


Figura 10. Relaciones de las entidades de la base de datos.

4.3 Descripción de los objetos de la base de datos

La base de datos construida posee los siguientes objetos:

- Tablas
- Consultas
- Formas
- Reportes

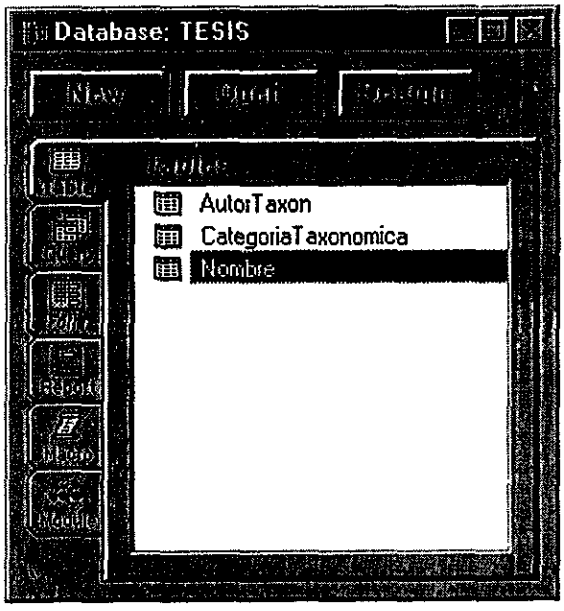


Figura 11. Vista de la ventana de Access que muestra las entidades o tablas de la base de datos.

4.3.1 Entidades

CATEGORIATAXONOMICA:

Contiene los nombres de las categorías taxonómicas desde Reino hasta las categorías infraespecíficas. Los campos que contiene son:

- IdCategoriaTaxonomica

  Clave primaria de tipo numérico contador (*counter*, asigna automáticamente un número consecutivo) de la entidad CATEGORIATAXONOMICA, funciona como un identificador único de cada registro.

- NombreCategoría

  Contiene el nombre de cada categoría taxonómica.

- Nivel

  Numeración consecutiva que indica jerarquía: las categorías asignadas con números enteros se consideran categorías obligatorias (comúnmente llamadas lineanas, por ejemplo: reino, división, clase), y aquellas con números fraccionarios se denominaron no obligatorias (subreino, subdivisión, superclase, subespecie).

- NOMBRE:

Es la entidad que contiene los nombres de los taxones. Está compuesta de nueve campos:

- IdNombre

  Clave primaria de tipo numérico contador de la entidad NOMBRE, funciona como un identificador único.

- IdAutorTaxon

  Clave foránea de la entidad AUTOR, relaciona cada taxón con su autoridad mediante un identificador.

- IdNombreAscendente

  Identificador (correspondiente al IdNombre) que establece una relación unaria de un taxón de la tabla NOMBRE con su taxón ascendente inmediato (categorías no obligatorias o intermedias).

- IdAscendenteObligatorio

  Identificador (correspondiente al IdNombre) que establece una relación unaria de un taxón de la entidad NOMBRE con su taxón ascendente obligatorio.

- Num_Filogenetico

  Nivel filogenético de tipo numérico que le corresponde según el autor de la clasificación. Con este es posible recuperar el árbol filogenético de los taxones.

- Nombre

  Es el nombre publicado del taxón.

- IdCategoriaTaxonomica

  Clave foránea de la tabla CATEGORIATAXONOMICA, relaciona el nombre del taxón con la categoría que le corresponde a través de su identificador.

- AnioCreacion

  Es el año de creación del nombre válidamente publicado.

- *SistemaDeClasificacion*

  Contiene la cita del sistema de clasificación, el nombre del autor y año.

Posteriormente se añadieron los campos Estatus y Fuente.

En esta entidad se representaron todas las clasificaciones seleccionadas. A continuación se ejemplifican dos clasificaciones, una para plantas y una para animales (cuadros 3 y 4).

Id Nombre	Id Autor Taxon	IdNombre Ascendente	Id Ascendente Obligatorio	NUM Filo GENético	Nombre	IdCategoría Taxonómica	Año creación	Sistema de clasificación
1	1	1	1	14603.1	PLANTAE	1	9999	Mabberley,
2	1	1	1	0.01	MAGNOLIOPHYT	2	9999	Mabberley,
3	1	2	2	0.02	MAGNOLIOPSIDA	4	9999	Mabberley,
4	1	3	3	0.021	MAGNOLIIDAE	5	9999	Mabberley,
6	1	4	3	0.03	Magnoliales	7	9999	Mabberley,
7	1	6	6	1	Winteraceae	9	9999	Mabberley,
8	1	6	6	2	Degeneriaceae	9	9999	Mabberley,
9	1	6	6	3	Himantandraceae	9	9999	Mabberley,
10	1	6	6	4	Eupomatiaceae	9	9999	Mabberley,
11	1	6	6	5	Austrobaileyaceae	9	9999	Mabberley,
12	1	6	6	6	Magnoliaceae	9	9999	Mabberley,
13	1	6	6	7	Lactoridaceae	9	9999	Mabberley,
14	1	6	6	8	Annonaceae	9	9999	Mabberley,
15	1	6	6	9	Myristacaceae	9	9999	Mabberley,
16	1	6	6	10	Canellaceae	9	9999	Mabberley,
17	1	4	3	0.03	Laurales	7	9999	Mabberley,
18	1	17	17	11	Amborellaceae	9	9999	Mabberley,
19	1	17	17	12	Trimeniaceae	9	9999	Mabberley,
20	1	17	17	13	Monimiaceae	9	9999	Mabberley,
21	1	17	17	14	Gomortegaceae	9	9999	Mabberley,
22	1	17	17	15	Calycanthaceae	9	9999	Mabberley,
23	1	17	17	16	Idiospermaceae	9	9999	Mabberley,
24	1	17	17	17	Lauraceae	9	9999	Mabberley,
25	1	17	17	18	Hernandiaceae	9	9999	Mabberley,

Cuadro 3. Ejemplo de los datos de plantas vasculares contenidos en la entidad NOMBRE. El "9999" es una convención para designar datos no disponibles.

Id Nombre	Id Autor Taxon	IdNombre Ascendente	Id Ascendente Obligatorio	NUM Filo GENético	Nombre	IdCategoría Taxonómica	Año creación	Sistema de clasificación
13802	1	13802	13802	0.1	ANIMALIA	1	9999	Nelson, 1993
13803	1	13802	13802	0.01	CHORDATA	24	9999	Nelson, 1993
13804	1	13803	13803	0.011	HEMICHORDATA	25	9999	Nelson, 1993
13805	1	13804	13803	0.02	PTEROBRANCHIA	4	9999	Nelson, 1993
13806	1	13804	13803	0.02	ENTEROPNEUSTA	4	9999	Nelson, 1993
13807	1	13803	13803	0.011	UROCHORDATA	25	9999	Nelson, 1993
13808	1	13803	13803	0.011	CEPHALOCHORDATA	25	9999	Nelson, 1993
13809	1	13803	13803	0.011	VERTEBRATA	25	9999	Nelson, 1993
13810	1	13809	13803	0.02	AGNATHA	4	9999	Nelson, 1993
13811	1	13810	13810	0.03	Petromyzontia	7	9999	Nelson, 1993
13812	1	13810	13810	0.03	Myxinoidea	7	9999	Nelson, 1993
13813	1	13809	13803	0.02	ELASMOBRANCHIOMORPHI	4	9999	Nelson, 1993
13814	1	13813	13813	0.021	CHONDRICHTHYES	5	9999	Nelson, 1993
13815	1	13814	13813	0.022	ELASMOBRANCHII	26	9999	Nelson, 1993
13816	1	13815	13813	0.03	Selachii	7	9999	Nelson, 1993
13817	1	13815	13813	0.03	Batoidea	7	9999	Nelson, 1993
13818	1	13814	13813	0.022	HOLOCEPHALI	26	9999	Nelson, 1993
13819	1	13818	13813	0.03	Chimaeriformes	7	9999	Nelson, 1993
13820	1	13809	13803	0.02	OSTEICHTHYES	4	9999	Nelson, 1993

Cuadro 4. Ejemplo de los datos de vertebrados contenidos en la entidad NOMBRE. El "9999" es una convención para designar datos no disponibles.

AUTORTAXON:

Contiene los nombres de todos los autores de cada género, especie o categoría *infraespecífica de la entidad* NOMBRE (cuadro 5).

IdAutorTaxon

Clave primaria de tipo numérico contador de la entidad AUTORTAXON, funciona como un identificador único (Contador o *Counter*).

Autor

Nombre del autor o grupo de autores.

IdAutor	Autor
1	(A Gray) W.F.Hillebr.
2	(A.Berger) Britton & Rose
3	(A.Berger) H.Ohba
4	(A.Berger) Riccob.
5	(A.Chev.) A.Chev. & Beille
6	(A.DC.) Boiss.
7	(A.DC.) Daniell
8	(A.DC.) DC.
9	(A.DC.) Hook.f.
10	(A.DC.) Mez
11	(A.DC.) Wight & Arn.
12	(A.Gray) A.Gray
13	(A.Gray) A.Nelson

Cuadro 5. Ejemplo de los datos contenidos en la entidad AUTORTAXON.

4.3.2 Consultas

Para validar los datos se elaboraron consultas de tal manera que se detectaran posibles errores y así se eliminaron autores y taxones duplicados, especies sin fecha de descripción y sin autor, y taxones pertenecientes a *más de un ascendente para una* misma clasificación.

Con ayuda de las consultas de Access se pueden crear al menos cuatro tipos de consultas (McGee and Boyce, 1993):

a) Consulta de selección: crea consultas acerca de preguntas específicas que se hace un usuario. Las principales consultas que se realizaron a la base son las siguientes:

a.1) Obtención de una lista de referencia de especies incluyendo o no autores y años de las mismas.

a.2) Recuperación de toda la clasificación desde reino hasta especie con todas las categorías ascendentes en un orden establecido, por ejemplo, alfabéticamente.

a.3) Recuperación de toda la clasificación desde reino hasta especie con todas las categorías ascendentes obligatorias o lineanas (figura 12).

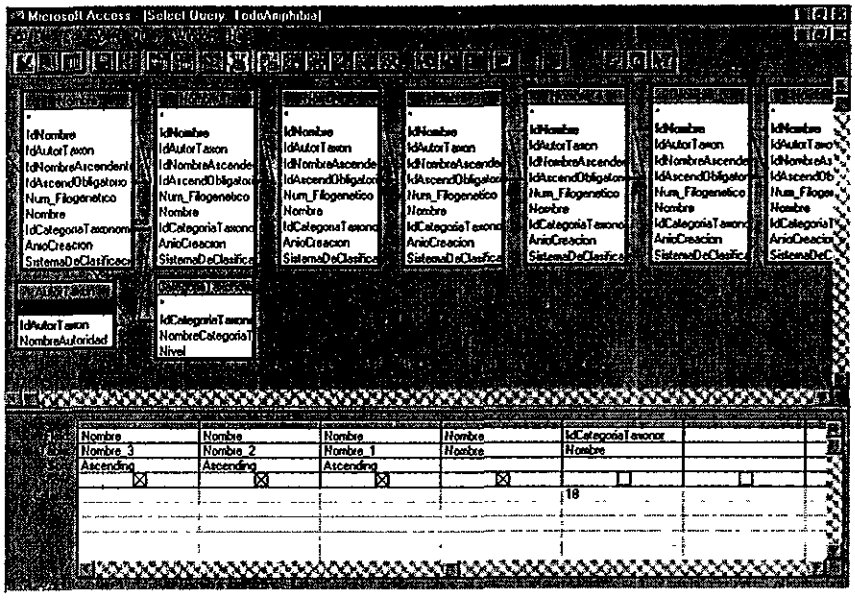


Figura 12. Ejemplo de consulta para obtener la clasificación completa de anfibios con las categorías obligatorias.

a.4) Recuperación de toda la clasificación filogenética desde reino hasta especie con todas las categorías ascendentes.

a.5) Recuperación de todos los nombres de autores incluyendo años en los que han publicado (por ejemplo: obtener un intervalo de tiempo en que un autor produjo muchas descripciones de especies).

b) Consulta para creación de tablas: crea una nueva tabla a partir de datos seleccionados de varias tablas. Cualquier tipo de consulta de selección puede ser convertida en tabla incorporándola a la base de datos y realizar más consultas a partir de esta. Por ejemplo, crear subtablas de un sistema de clasificación con base en un criterio específico.

c) Consulta para actualizar: cambia conjuntos de datos en las tablas. Cuando existen modificaciones en el sistema de clasificación de cierto taxón, es posible actualizar la información de los registros que contienen cierto nombre ascendente de manera automática para todos los taxones que se involucren en el cambio (figura 13).

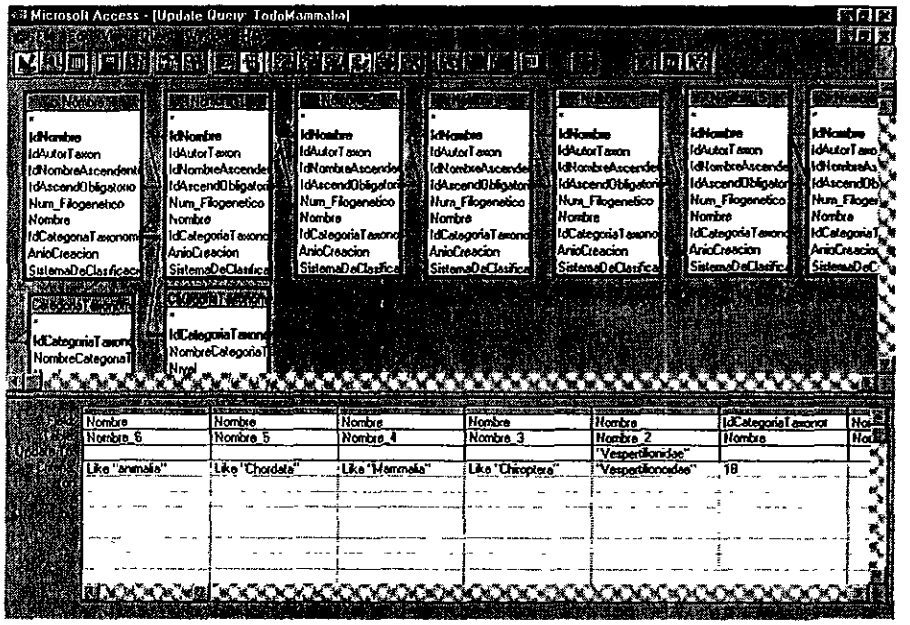


Figura 13. Ejemplo de consulta para actualizar registros en la clasificación de mamíferos.

d) Consulta de añadir datos: adiciona registros a las tablas. Permite incorporar a una tabla de una base de datos existente un grupo de registros de un taxón determinado pertenecientes a una base de datos totalmente ajena a la primera.

e) Consulta de eliminación: remueve registros de las tablas. Elimina registros de taxones que hayan sido ingresados equivocadamente a la base de datos.

4.3.3 Formularios

Este tipo de herramientas permitió que se llevara a cabo la captura de los datos de una forma más ordenada ya que al disponer de pantallas de captura los datos se ingresaron de la forma más conveniente eliminando algunos errores de tecleo.

a) Formulario para capturar autores de taxones (figura 14).

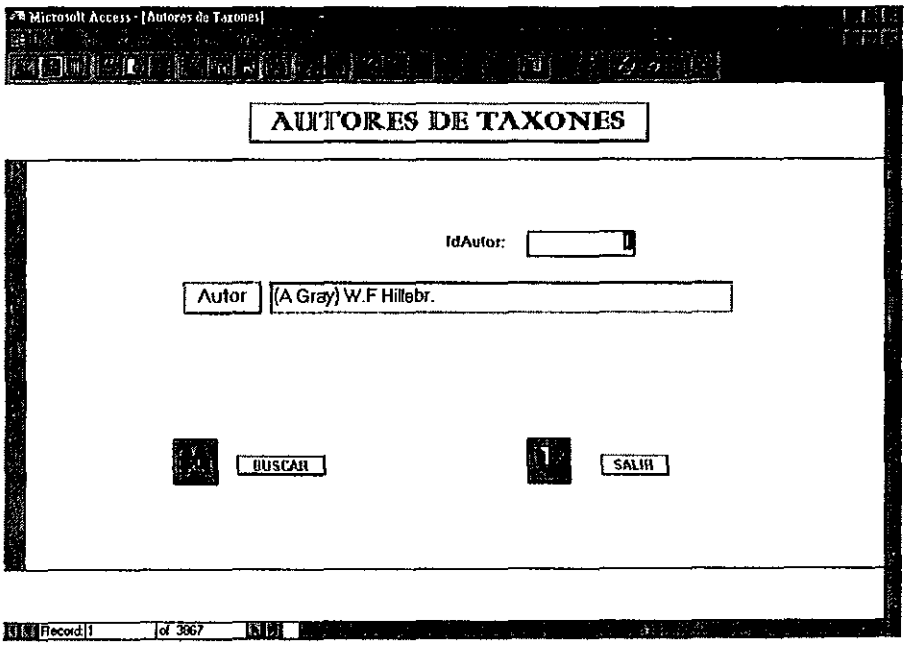


Figura 14. Ejemplo de formulario para ingresar autores nuevos.

b) Formulario para captura de nombres y ascendentes nuevos (figura 15).

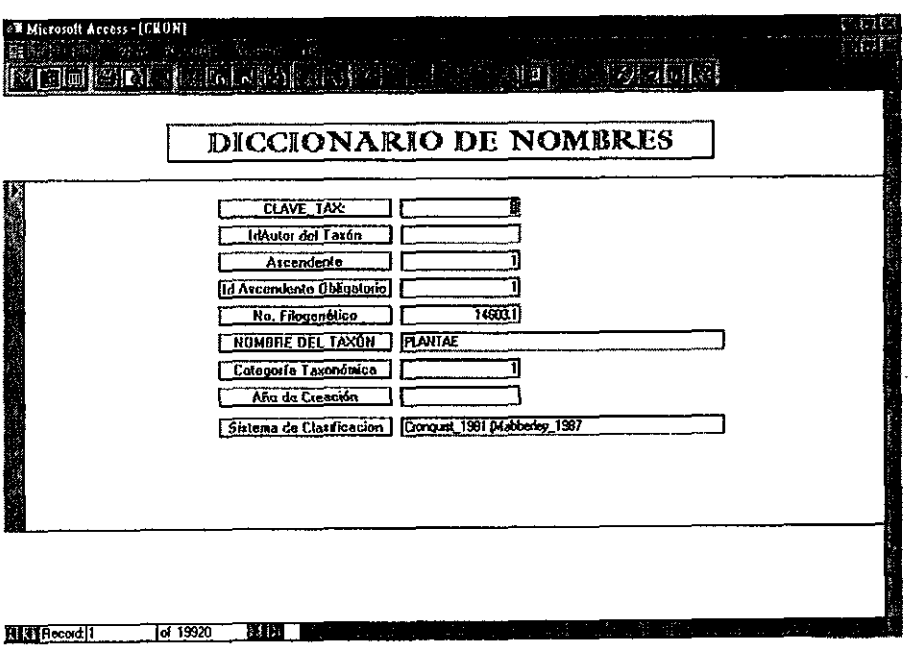


Figura 15. Ejemplo de formulario para captura de nombres y ascendentes de plantas vasculares.

c) Formulario para consulta y depuración de los sistemas de clasificación (figura 16).

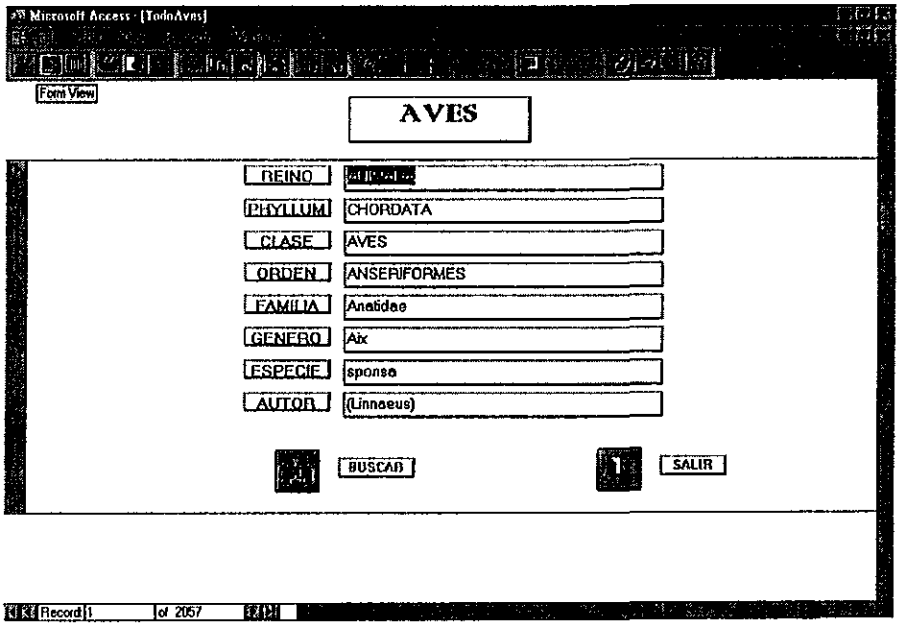


Figura 16. Ejemplo de formulario para revisar y depurar la clasificación de aves.

4.3.4 Reportes

Se pueden generar reportes o informes a partir de cualquier tabla o consulta prácticamente para cualquier atributo o conjunto de ellos; pudiendo utilizar ciertos criterios, estos reportes se puede generar para cualquier taxón, independientemente del nivel en el que se encuentre, en los cuales se pueden incluir gráficos, estadísticas, etc. En las siguientes figuras (17 y 18) se muestran el diseño y la presentación de un reporte elaborado para anfibios cuya fuente es la consulta de la figura 12.

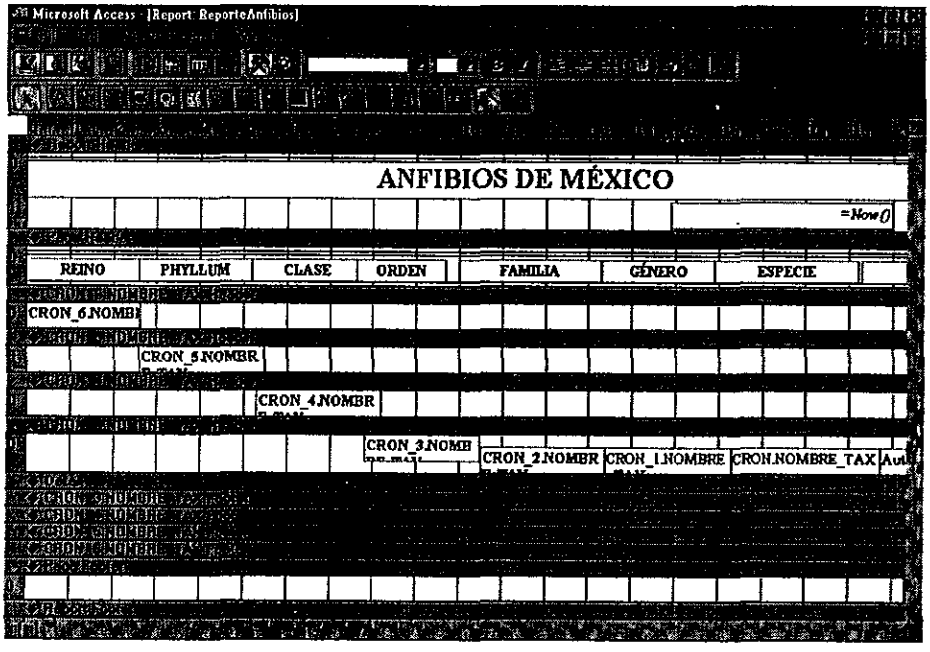


Figura 17. Ejemplo del diseño de un reporte para el sistema de clasificación de anfibios.

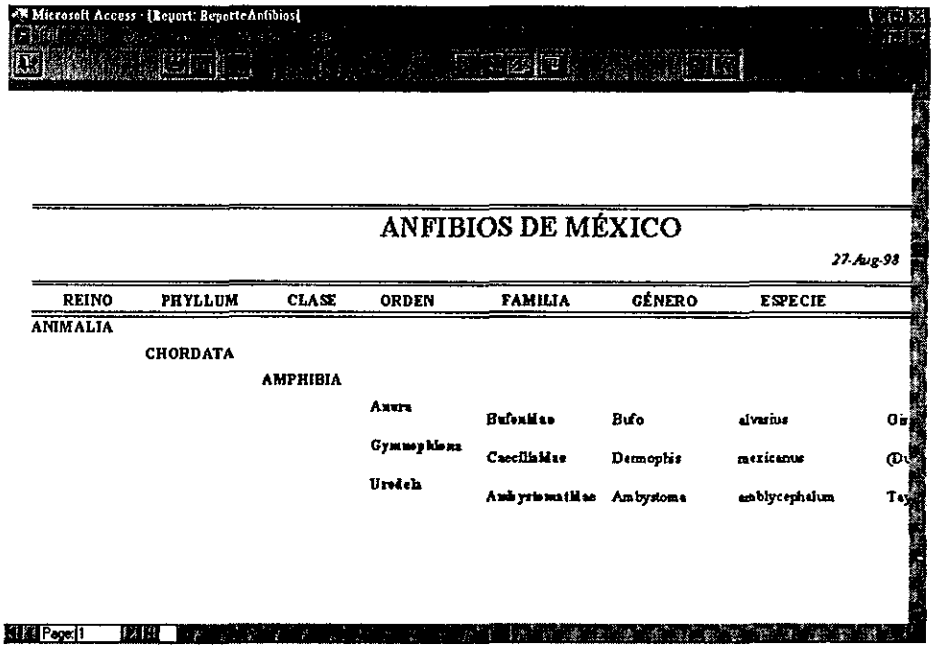


Figura 18. Ejemplo de la presentación preliminar de un reporte para el sistema de clasificación de anfibios.

4.4 Descripción de los nombres de taxones

Se capturaron alrededor de 35,180 nombres de taxones de plantas vasculares y vertebrados de México, para más de 30 diferentes categorías taxonómicas, y cerca de 3,900 autores.

Se capturaron 3,867 nombres de autores, incluyendo autores únicos y combinaciones ordenadas de dos o más autores

Las categorías taxonómicas involucradas son 34 (cuadro 6).

IdCategoriaTaxonomica	NombreCategoria	Nivel
1	reino	1
29	subreino	1.1
30	superphyllum	1.9
2	division	2
24	phyllum	2
3	subdivision	2.1
25	subphyllum	2.1
27	superclase	2.7
28	grado	2.9
4	clase	3
5	subclase	3.1
26	infraclase	3.3
6	superorden	3.9
7	orden	4
8	suborden	4.1
31	infraorden	4.3
32	superfamilia	4.9
9	familia	5
10	subfamilia	5.1
33	supertribu	5.3
11	tribu	5.5
12	subtribu	5.7
13	genero	6
14	subgenero	6.1
15	seccion	6.3
16	subseccion	6.5
17	serie	6.7
34	subserie	6.9
18	especie	7
19	subespecie	7.1
20	variedad	7.3
21	subvariedad	7.5
22	forma	7.7
23	subforma	7.9

Cuadro 6. Categorías taxonómicas en la base de datos. Las categorías designadas por enteros en el campo nivel comunes a todos los taxones.

En el siguiente cuadro se muestran cuántas categorías se capturaron por cada uno de los principales sistemas de clasificación de la base de datos:

CATEGORÍA	TOTAL
Reino	2
Subreino	1
Superphyllum	4
División	13
Phyllum	2
Subdivisión	0
Subphyllum	5
Superclase	2
Grado	2
Clase	34
Subclase	27
Infraclase	4
Superorden	19
Orden	267
Suborden	115
Infraorden	3
Superfamilia	10
Familia	1630
Subfamilia	152
Supertribu	0
Tribu	51
Subtribu	0
Género	27336
Subgénero	10
Sección	0
Subsección	0
Serie	0
Subserie	0
Especie	4465
Subespecie	1030
Variedad	0
Subvariedad	0
Forma	0
Subforma	0
Total	***35186***

Cuadro 7. Categorías taxonómicas obligatorias y no obligatorias consideradas en la base de datos.

A continuación se muestran los diferentes sistemas de clasificación que se utilizaron para realizar los diccionarios:

GRUPO	SISTEMA
PLANTAS VASCULARES	
Pteridofitas	R.C. Moran & R.Riba (1985)
Gimnospermas	Bold, Alexopoulos y Delevoryas (1980)
Angiospermas	C.Bentham & J.D.Hooker (1862-1883) modificado por Brummitt (1993) A. Cronquist (1981) modificado por Mabberley (1987)
MAMÍFEROS	
Mamíferos terrestres	G. Ceballos y H. Arita (1996)
	J. Ramírez-Pulido, et al. (1996) modificado de Wilson y Reeder (1993)
Mamíferos marinos	M. Salinas y P. Ladrón de Guevara modificada de Barnes, et al., (1985) y King (1983)
HERPETOFAUNA	Flores-Villela (1993); modificado de Dudle (1989), Frost (1985), King & Burke (1989) y Dubois (1985)
PECES DULCEACUÍCOLAS	H. Espinosa, et al., (1993) modificado de Nelson (1984)
AVES	Check List of North American Birds by The Committee on Classification and Nomenclature of the American Ornithologist's Union (1983)

Cuadro 8. Sistemas de clasificación utilizados para cada taxón.

En los siguientes cuadros se presenta un resumen de los sistemas de clasificación por taxón y el número de especies y otras categorías para cada uno de ellos (cuadros 9 y 10):

GRUPOS	SISTEMAS DE CLASIFICACIÓN	No. DE ESPECIES/SUBESPECIES
Pteridofitas, Gimnospermas y Angiospermas	Brummitt (1993)	12,956 (géneros)
	Cronquist (1981)	12,965 (géneros)
Mamíferos	G. Ceballos y H. Arita (1996)	507
	Ramírez-Pulido, et al (1996)	491/1028
Anfibios y Reptiles	Flores-Villela (1993)	995
Peces Dulceacuícolas	Espinosa, et al. (1993)	403/2
Aves	AOU (1993)	2099

Cuadro 9. Número de especies/subespecies por taxón y sistema de clasificación.

GRUPOS		FAMILIAS	GENEROS	ESPECIES	SUBESPECIES	TOTAL
Plantas Vasculares	Brummitt (1993)	497	12, 956	0	0	13,453
	Cronquist (1981)	428	12, 965	0	0	13,393
Mamíferos	Ceballos (1996)	44	188	507	0	739
	Ramírez-Pulido (1996)	46	187	491	1028	1752
Anfibios (Flores-Villela, 1993)		14	45	290	0	349
Reptiles (Flores-Villela, 1993)		37	154	705	0	896
Peces Dulceacuícolas (Espinosa et al., 1993)		482	110	403	2	997
Aves (AOU, 1993)		93	756	2099	0	2948
TOTAL		***1641***	***27361***	***4495***	***1030***	***34527***

Cuadro 10. Número de familias, géneros y especies totales para cada taxón.

V. ANÁLISIS DE RESULTADOS

5.1 Colecciones biológicas y sistemas de clasificación

Las especies son las unidades con las cuales medimos la biodiversidad, son esenciales para formar unidades taxonómicas claras y estables, y son el soporte para desarrollar bases de datos sobre biodiversidad y trabajar en programas cooperativos nacionales e internacionales de análisis, planeación, manejo y conservación. El conocimiento de los taxones, sus interrelaciones, clasificación, evolución, distribuciones pasadas y recientes, y el entendimiento de su papel en el ecosistema, es información que en ocasiones está disponible sólo en colecciones biológicas y bases de datos. Al presentar cambios continuos e inconsistencias en la nomenclatura de los taxones, el diseño de bases de datos es una tarea *ad hoc* para compilar información que posteriormente puede ser usada para elaborar programas de conservación (Nielsen & West, 1994; Green, 1994).

En México, según datos de la CONABIO, existen cerca de 60 sedes institucionales con cerca de 180 colecciones científicas, las cuales albergan casi 10,000,000 ejemplares (cuadro 11). De las colecciones botánicas, el 61.11% tiene ejemplares de plantas vasculares. En el caso de las colecciones zoológicas, el 56.3% incluyen vertebrados, de las cuales 12.41% son peces, 11.76% reptiles, 10.45% anfibios, 9.8% aves y 9.15% mamíferos (Llorente y Koleff, 1997).

Área	Número de colecciones institucionales	Número de ejemplares
Botánica	72	2,931,640
Zoología	102	6,801,800
Microbiología	5	14,630
Total	***179***	***9,748,070***

Cuadro 11. Datos generales de las instituciones registradas en CONABIO que albergan colecciones científicas, según Llorente y Koleff (1997).

La CONABIO, como organismo coordinador y promotor de las investigaciones científicas sobre los recursos biológicos, debe conocer a los proveedores de la información básica: los taxónomos y los biogeógrafos, quienes utilizan el sistema de clasificación y la nomenclatura que permite reconocer a las especies, nombrarlas y describirlas (Llorente y Koleff, 1997). De ahí proviene el interés de desarrollar el sistema BIOTICA, el cual, como parte fundamental de su estructura incluye un bloque nomenclatural a través de árboles que representen los sistemas de clasificación taxonómica. Para dichos árboles se consideraron en este trabajo, para cierots taxones más de un sistema de clasificación, manejados como ejes y alternativos.

Los sistemas de clasificación eje para plantas vasculares y mamíferos no resultaban suficientes ya que el número de taxones totales para el país no se encontraba bien definido en un sólo sistema, es decir, no existe un consenso real sobre cuál sistema usar entre los investigadores que tienen a su cargo las principales colecciones biológicas. El grupo de las plantas vasculares es tan extenso que incluso no existe un listado de referencia nacional ni un sistema de clasificación a nivel específico que abarque todas las familias, y para los mamíferos, continuamente existen cambios a niveles infraespecífico.

Aunque la incorporación de sistemas de clasificación alternativos fue realizada en primera instancia como un complemento al sistema de clasificación eje, es necesario que sean incluidas relaciones de equivalencias entre los nombres de los taxones de ambos sistemas de clasificación. La principal dificultad de llevar a cabo esta tarea, es el uso de los nombres a los diferentes taxones. Según Berendsohn (1998), la circunscripción del taxón no debe confundirse con el nombre. La circunscripción está definida por el taxónomo, usualmente en un tratamiento monográfico, basados en tipos. Si bien el concepto de tipo es de dar estabilidad y evitar que se incrementen los nombres, no hay una clave única del concepto del taxón. La solución potencial que se propone es unir conceptos y nombres y tener un "*potential taxa*" (taxón potencial), es decir, nombres con información de circunscripción ligados a ellos con una referencia bibliográfica. Dado que hay sinónimos nomenclaturales y conceptuales, la implementación del taxón potencial puede proveer un archivo de autoridad (al mezclar nombres y conceptos), ligando información no ambigua al diccionario (figura 19).

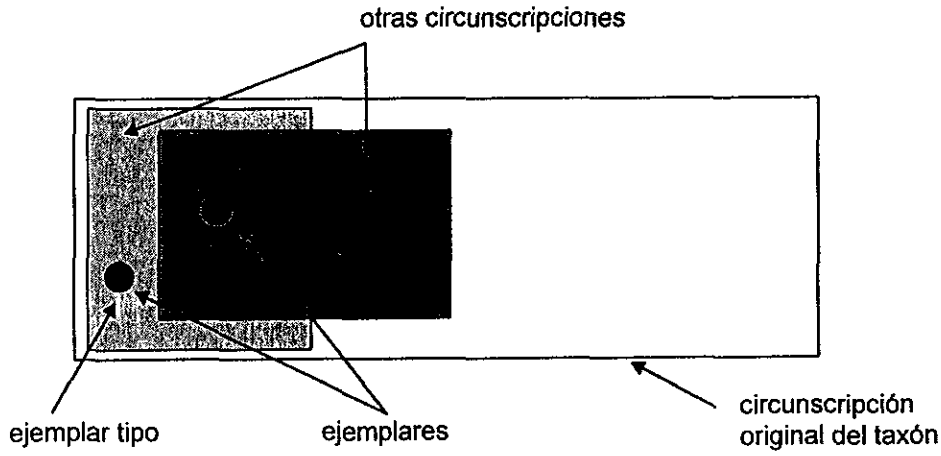


Figura 19. Cambios en la circunscripción de un taxón, según Berendsohn (1998).

En el caso de los peces dulceacuícolas, herpetofauna y aves, sólo existe un tratamiento taxonómico completo en los primeros grupos y uno reconocido internacionalmente para las aves, por lo que no fue necesario incluir sistemas de clasificación alternativos.

El siguiente paso para conformar el diccionario de nombres aceptados de un taxón, sería la incorporación de los nombres sinónimos a cada uno de ellos, mediante una entidad a la que llamaremos SINONIMIANOMENCLATURAL, la cual funcionará como una entidad relacional, formada a través de una relación muchos-a-muchos con la entidad NOMBRE.

5.2 Modelo Entidad-Relación: tratamiento de los datos nomenclaturales

El desarrollo de un modelo de bases de datos que represente fielmente un sistema de clasificación taxonómica jerárquica es realmente complejo.

Consideremos primeramente el modelo como una sola entidad (*flat-file*) a la que llamaremos NOMBRE, en la cual la jerarquía se encuentra vinculada de la manera más simple, es decir, dicha entidad contendría todas las categorías en un mismo registro. El modelo permite una lectura rápida directamente a través de la tabla, pero la captura no resulta tan eficiente, ya que se requiere teclear la información repetidas veces y esto genera errores. Cuadro 12.

Reino	Phyllum	Clase	Orden	Familia	Género	Especie
ANIMALIA	CHORDATA	REPTILIA	Crocodilia	Alligatoridae	Caiman	crocodylus
ANIMALIA	CHORDATA	REPTILIA	Crocodilia	Crocodylidae	Crocodylus	acutus
ANIMALIA	CHORDATA	REPTILIA	Crocodilia	Crocodylidae	Crocodylus	moreleti
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	bogerti
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	chiszari
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	deppei
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	fuscolabialis
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	graminea
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	kalaina
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	lythrochila
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	matudai
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	mitchelli
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	mixteca
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	oaxacae
ANIMALIA	CHORDATA	REPTILIA	Squamata	ANGUIDAE	Abronia	ochoterenai
ANIMALIA	CHORDATA	REPTILIA	Squamata	Anguidae	Abronia	omelasi

Cuadro 12.Tabla única para una clasificación de reptiles.

En el caso de que quisiéramos representar cada jerarquía taxonómica por medio de una entidad, obtendríamos tantas entidades como jerarquías tuviera el sistema. En este trabajo, tendríamos al menos siete entidades relacionadas que corresponden a las categorías obligatorias (figura 20).

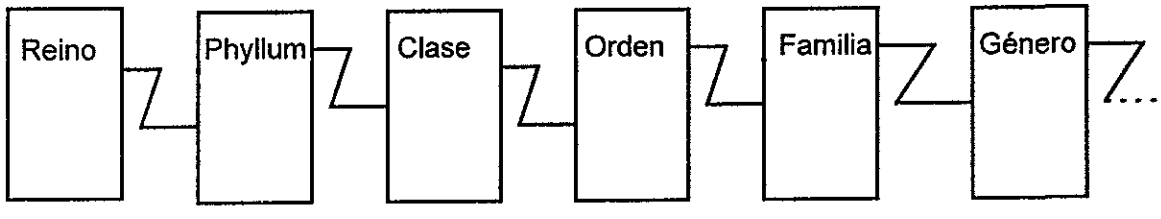


Figura 20. Modelo de datos donde cada entidad corresponde a una jerarquía taxonómica.

La principal desventaja es cuando un sistema de clasificación es modificado adicionando un nivel taxonómico que anteriormente no existía, tiene que insertarse una entidad intermedia y modificarse el modelo de datos (figura 21).

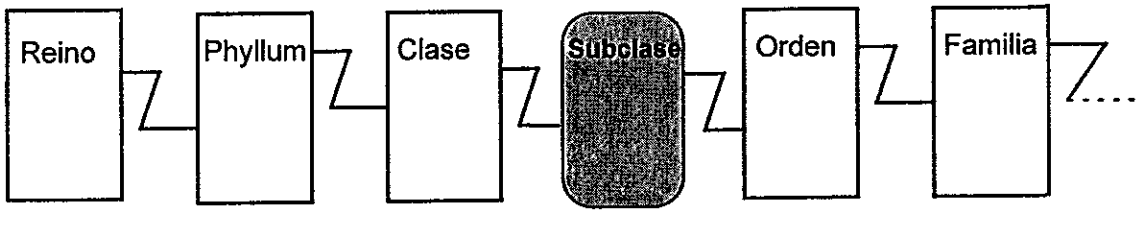


Figura 21. Modificación al modelo de datos donde se ingresa una categoría internedia.

Para realizar una base de datos más eficiente, la solución sería contar con el mínimo de entidades y al mismo tiempo evitar capturar información repetidamente. Un modelo de datos que contenga relaciones unarias ofrece dichas ventajas y además permite obtener toda la información como una tabla única por medio de una consulta.

El diseño se puede visualizar como un árbol inf-reticulado (figura 22):

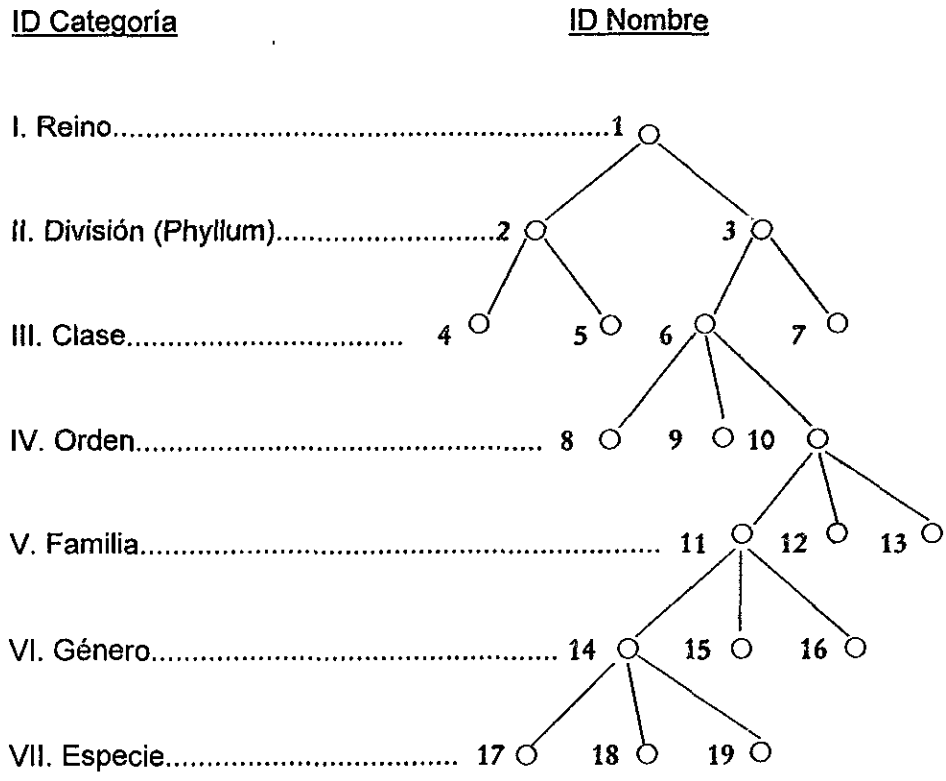


Figura 22. Esquema del árbol de categorías y nombres de taxones que puede representarse como una entidad con relaciones unarias.

donde cada nodo corresponde a un nombre en una categoría taxonómica específica y cada línea a una relación ascendente.

Sigamos entonces las bases de la metodología de modelado presentada por Zellweger & Allkin (1993):

El primer paso del diseño de una base de datos es generar un modelo conceptual que identifique al objeto bajo estudio y sus propiedades. El modelo puede también definir las relaciones entre esos objetos. Un taxón (tal como variedad, especie, género o familia) es una entidad considerada como una "realidad" biológica y representa una colección de poblaciones o especímenes individuales. Un nombre, cualquier binomio, trinomio o uninomio en latín, es una etiqueta usada como sinónimo o nombre aceptado, para referirse a una de esas entidades biológicas o abstractas (Zellweger & Allkin, 1993).

El segundo paso es usar diagramas de datos entidad-relación para modelar esta "realidad". Los taxones individuales están relacionados uno a otro dentro de un sistema jerárquico. Existe una serie de relaciones uno-a-muchos entre pares de taxones: cada taxón incluye varios taxones al siguiente nivel inferior y él mismo corresponde a sólo un taxón al siguiente nivel superior. Entre los nombres se pueden establecer relaciones uno-a-muchos entre ellos mismos: ciertas variantes ortográficas pueden ser ligadas a una ortografía "correcta", sinónimos homotípicos pueden ser ligados a un basónimo, y sinónimos heterotípicos pueden ser ligados a un nombre aceptado (Zellwege & Allkin, 1993).

Para los fines de este trabajo cada taxón está representado por un registro de la entidad NOMBRE el cual tiene a su vez un cierto número de taxones tanto por arriba de él (categorías superiores) como por abajo (categorías inferiores), es decir, hay una jerarquía entre ellos. En la entidad NOMBRE se encuentra representada una relación uno-a-muchos de manera unaria, es decir, la entidad siempre hace referencia a sí misma en el mismo campo (figura 23). Así, un taxón ***X*** en un nivel determinado tiene por debajo de él ***n*** taxones (por ejemplo, una familia podría poseer tres géneros, un órden puede estar formado por cinco familias, etc.), pero nunca ese taxón ***X*** puede estar contenido en más de un taxón superior (una familia en dos órdenes, un órden en dos clases, etc.) . Las relaciones unarias se establecen mediante el identificador único y la repetición de éste como identificadores que hacen referencia a los nombres de sus taxones superiores (IDNOMBRE y IDNOMBREASCENDENTE y IDASCENDENTEOBLIGATORIO, respectivamente).

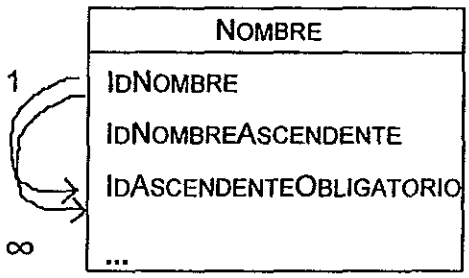


Figura 23. Esquema de la entidad NOMBRE y sus relaciones unarias.

En el modelo se han establecido las relaciones entre las entidades NOMBRE y CATEGORIATAXONOMICA y NOMBRE y AUTORTAXON las cuales también son relaciones uno-a-muchos (1 → ∞). Muchos nombres de taxones pueden pertenecer a una categoría (por ejemplo, Laurales y Magnoliales son órdenes), sin embargo, un taxón no puede pertenecer a más de una categoría (Magnoliales únicamente es órden, y no podría ser órden y clase a la vez). Igualmente, un autor o una combinación única de autores describió más de un nombre de taxón (por ejemplo, Linneo nombró a cientos de especies), y ningún taxón aceptado ha sido descrito por más de un autor o una combinación de autores. En ocasiones, los homónimos (nombres de taxones idénticos) pueden tener el mismo autor, sin embargo, el año de creación del nombre es el elemento que en este caso identifica a cada uno como registros diferentes (figura 24).

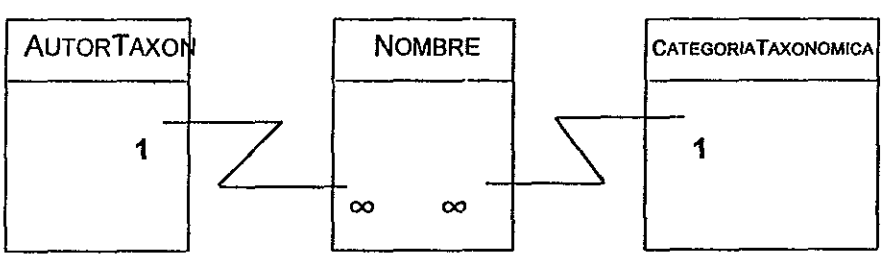


Figura 24. Esquema de las entidades NOMBRE, AUTORTAXÓN Y CATEGORIATAXONOMICA con sus relaciones.

El uso de diagramas entidad-relación permite la comparación y exploración de diferentes aproximaciones para visualizar un sistema de información en un campo particular. Se pueden producir modelos alternativos que independiente y subsecuentemente de las limitantes de la implementación, pueden ser usados como herramientas de comunicación en el proceso de diseño. Un diagrama entidad-relación da una visión clara y sin ambigüedades (Zellweger & Allkin, 1993).

Para ser utilizados, los datos deben confrontarse con estándares, los cuales juegan un papel crucial en la actividad coordinativa (Green, 1994). Los estándares son mecanismos usados en la recolección, manejo e intercambio de datos, los cuales han de ser adoptados. Los beneficios de usar estándares son (WCMC, 1997c):

- obtención de consistencia , consolidación y comunicación
- incremento en la eficiencia del conjunto de datos
- menores costros al realizar transacciones de intercambio de datos
- los datos pueden ser utilizados en diversas aplicaciones

El desarrollo de estándares generalmente requiere un esfuerzo interdisciplinario para revisar, formalizar y publicar los estándares, además deben establecerse procesos de revisión de los mismos (WCMC, 1997c).

Existen dos clases principales de estándares según Green (1994):

Atributos estándares.- se refieren a la información que es esencial para todos los datos; alguna información puede ser deseable pero no esencial.

En este caso particular, la información esencial para reconocer un nombre de un taxón es primeramente que debe pertenecer a una clasificación taxonómica y encontrarse válidamente publicado, debe además corresponder a un cierto nivel taxonómico dentro de una jerarquía y ser descrito por una autoridad en una fecha determinada.

Estándares de control de calidad.- proporcionan indicadores de validación y precisión para los datos, como referencias a glosarios o autoridades usadas para los nombres y códigos que incluyan las clases de error que han sido llevadas a cabo en el ingreso de los datos.

Un diccionario de nombres de taxones se considera entonces como un estándar de control de calidad, ya que al intersectar cierta información de otra base de datos con la base de datos estándar se pueden localizar fácilmente inconsistencias entre los datos.

En el SNIB, los archivos de autoridades funcionan como indicadores de validación al someter modelos de bases de datos diferentes a BIOTICA (por ejemplo, entidades "planas" y otros modelos relacionales) con información nomenclatural a una inspección de acuerdo con determinada clasificación. Así, podremos localizar y rechazar nombres de taxones que han sido ingresados a la base de datos que no corresponden a la clasificación mencionada, sin embargo, están correctamente escritos.

El actual uso de BIOTICA, con los diccionarios incorporados al modelo, reduce significativamente el número de inconsistencias presentes en las bases de datos hasta más de un 5%, sin embargo, en bases de datos con diferente modelo de datos, el número de errores se incrementa considerablemente, alcanzando alrededor del 20% (figura 25).

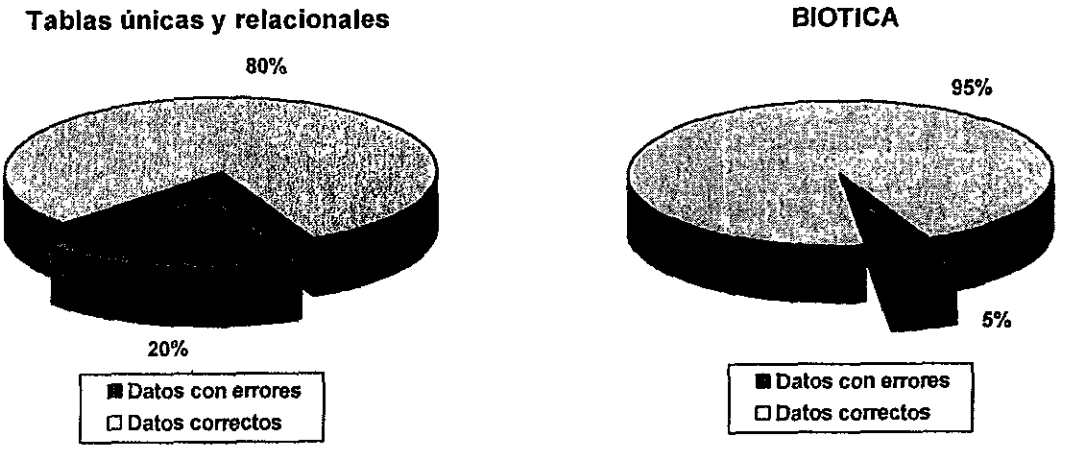


Figura 25. Comparación en el porcentaje de error al validar bases de datos con diccionarios y sin ellos. Fuente: Subdirección de Inventarios Bióticos, CONABIO.

Estos archivos también son indicadores de precisión puesto que permiten obtener mediante intersecciones qué nombres de taxones no corresponden a cierta clasificación debido a que han sido capturados erróneamente, es decir, nos pueden ayudar a detectar errores y hacer una base de datos más precisa en cuanto a su nomenclatura.

Sin embargo, la creación de estándares no es necesariamente sencilla, las dificultades para crear una lista de referencia de especies de manera global son de acuerdo con Bisby (1993) las siguientes:

1. La existencia de tratamientos taxonómicos alternativos (cuando dos revisiones taxonómicas están en conflicto, hay posiblemente dos sistemas estructurales diferentes).

Esta dificultad se puede minimizar empleando un sistema de clasificación eje y otros alternativos, como en el presente trabajo.

2. Integridad taxonómica (cada lista regional es cuidadosamente organizada y por lo tanto reconoce clasificaciones infraespecíficas muy particulares y diferentes).

Por ejemplo, en el caso de las plantas vasculares, un sistema que contenga todas las especies es casi imposible de obtener, debido a que cada pequeño grupo de taxones (por ejemplo, las especies de la familia Burseraceae) puede estar organizado en diferentes clasificaciones dependiendo del tratamiento al que ha sido sujeto el taxón por parte de los taxónomos.

3. Grupos pobremente conocidos. Para áreas grandes del mundo, y para muchos grupos, el tratamiento taxonómico es escaso, pobre o inexistente.

La realización de un diccionario de taxones para grupos pobremente conocidos es casi imposible de realizar por las implicaciones de elaboración que este requiere (ausencia de atributos estándares).

4. Tratamientos taxonómicos parciales. Las monografías y revisiones son raras para géneros más o menos grandes.

Igualmente en el caso de las plantas vasculares fue una limitante en el presente trabajo. Únicamente se incorporaron taxones hasta el nivel de género ya que no existen descripciones para grupos demasiado grandes (por ejemplo Compositae y Gramineae) ni una lista nacional que incorpore todas las especies.

5.3 Bases de datos nomenclaturales

Técnicas tales como el modelo entidad-relación que se ha discutido anteriormente, permite a los taxónomos definir los requerimientos de su sistema y hacer sus propias elecciones, lo cual reduce el impacto de las limitaciones por el *hardware* y *software* disponible y ser más independientes de los analistas de sistemas quienes pueden caer en el error de diseñar una base de datos pensando que conocen suficientemente el trabajo que hace un taxónomo. Las elecciones hechas en el modelo conceptual permiten definir con precisión los requerimientos para la implementación del sistema, lo que puede

eventualmente evaluar la selección de *software* y *hardware* compatible (Zellweger & Allkin, 1993).

La falta de financiamiento y otras presiones generalmente resultan en sistemas de información construidos precipitadamente. Para resolver este problema, se requiere desarrollar estrategias que excluyan nomenclatura inconsistente en los datos y eviten nuevas inconsistencias. Potencialmente las bases de datos pueden adoptar tres diferentes aproximaciones para reconciliar diferencias de los nombres dentro de los datos (Green, 1994):

-Desarrollar sistemas para coordinar cambios de nombres y asegurar la consistencia y la notificación de cambios en los nombres a usuarios del sistema.

-Crear identificadores para ligar taxones actuales con los nombres que han sido usados. Es decir, mantener un "registro histórico" de todos los cambios que se han realizado al nombre de un taxón.

-Vínculos a campos con datos de claves taxonómicas y descripciones para la identificación de especímenes.

5.4 Implementación del modelo de datos

5.4.1 Sección nomenclatural de BIOTICA

La implementación del modelo que se presenta para BIOTICA incluyó la elaboración de una interface gráfica (figura 26) para la captura de los datos curatoriales por parte de los usuarios. A continuación se muestra el empleo de tal interface para el ingreso, uso y actualización de lo diccionarios.

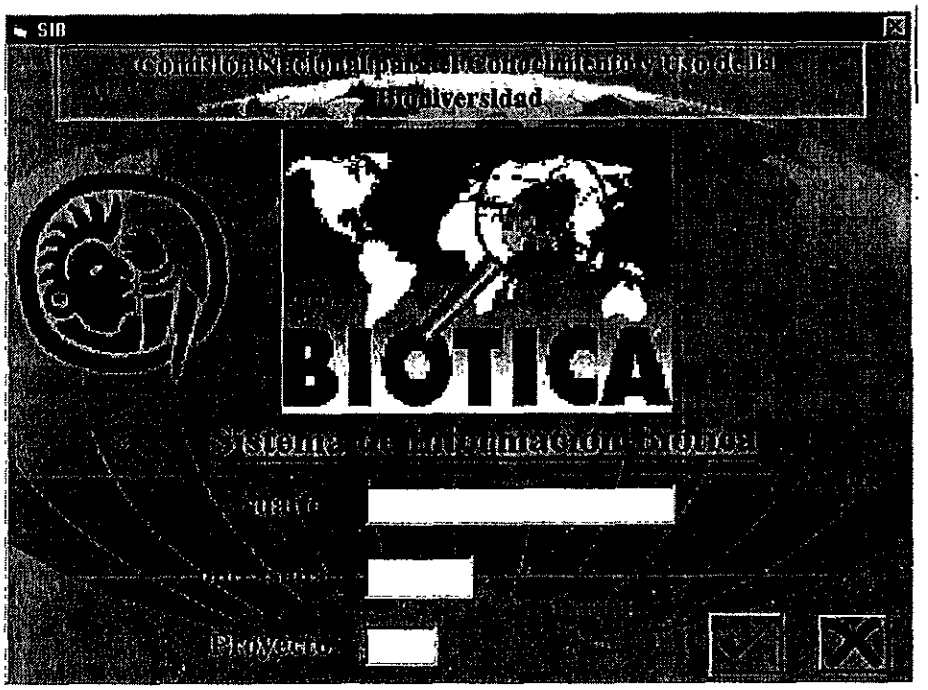


Figura 26. Ventana principal de acceso al Sistema de Información BIOTICA.

Al iniciar el programa y después de conectarse a la base de datos a utilizar, se encuentra disponible el menú principal (figura 27).

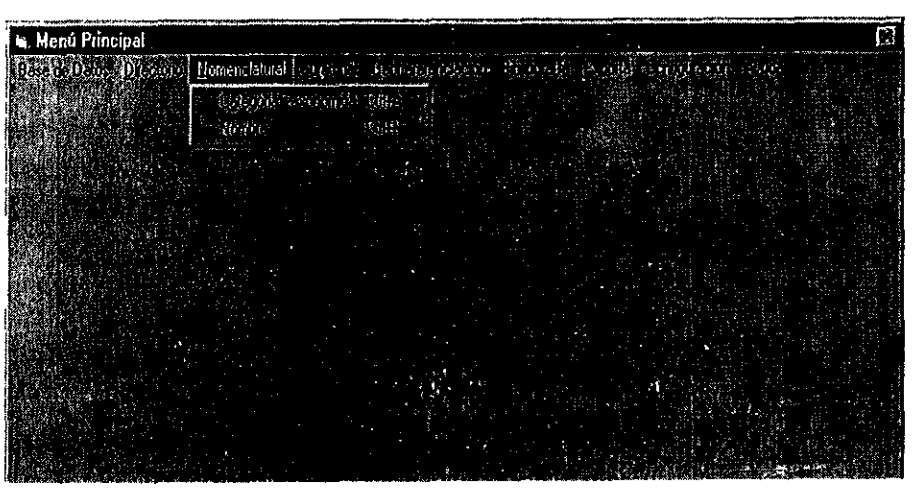


Figura 27. Menú de la sección nomenclatural.

La sección nomenclatural está compuesta de dos pantallas: Categoría Taxonómica y Nombre. Estas pantallas presentan una barra de menú y una de herramientas. A través de la barra de menú se tiene acceso al Catálogo de Autoridades (lista de autores del nombre del taxón) y a la Simbología (árbol taxonómico con los nombres de las categorías obligatorias y no obligatorias y su simbología).

La pantalla Catálogo de Categorías Taxonómicas contiene información sobre la jerarquias taxonómicas obligatorias y sus categorías intermedias, así como las infraespecíficas. Las primeras deben registrarse de manera obligada. Para expandir el árbol taxonómico es necesario dar doble "*click*" en cada nivel hasta llegar al nivel deseado. El signo de "+" significa que existen niveles intermedios; el signo de "-" significa lo contrario (figura 28).

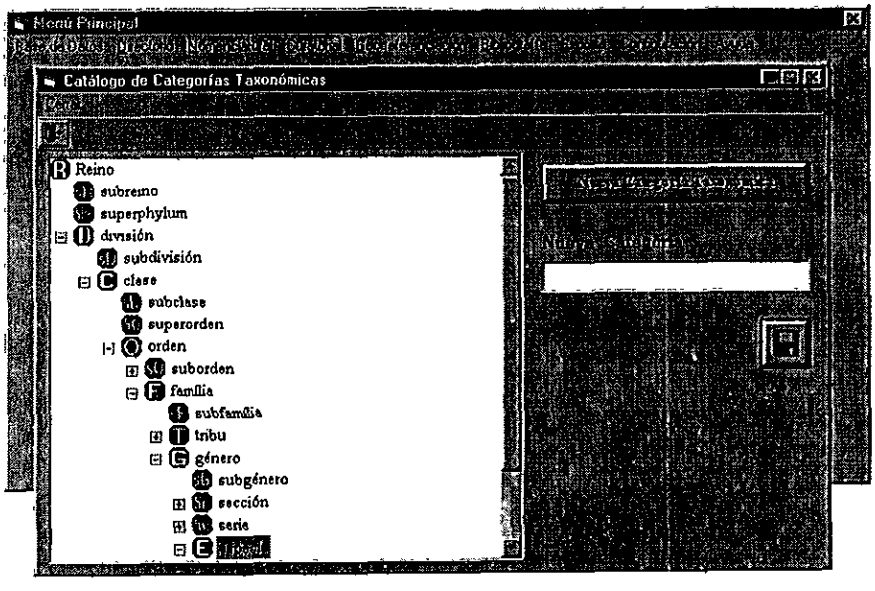


Figura 28. Pantalla del Catálogo de Categorías Taxonómicas.

La pantalla Captura y Modificación del Taxón contiene información de los nombres taxonómicos, estatus, sistema de clasificación, autor del taxón, número filogenético, sinonimias, basónimo/nombre original y registro de imagen.

Los nombres taxonómicos deben registrarse del nivel más general al particular, es decir, de phyllum o división hasta especie o subespecie.

El nombre de la autoridad o autor del taxón se selecciona a partir del Catálogo de Autoridades, dando un "*click*" en la barra del menú se selecciona el nombre correcto. Al dar doble "*click*" con el botón izquierdo en la parte gris del renglón, el autor se arrastra hasta la casilla correspondiente o se utiliza el icono de Traspaso de Datos (figura 29).

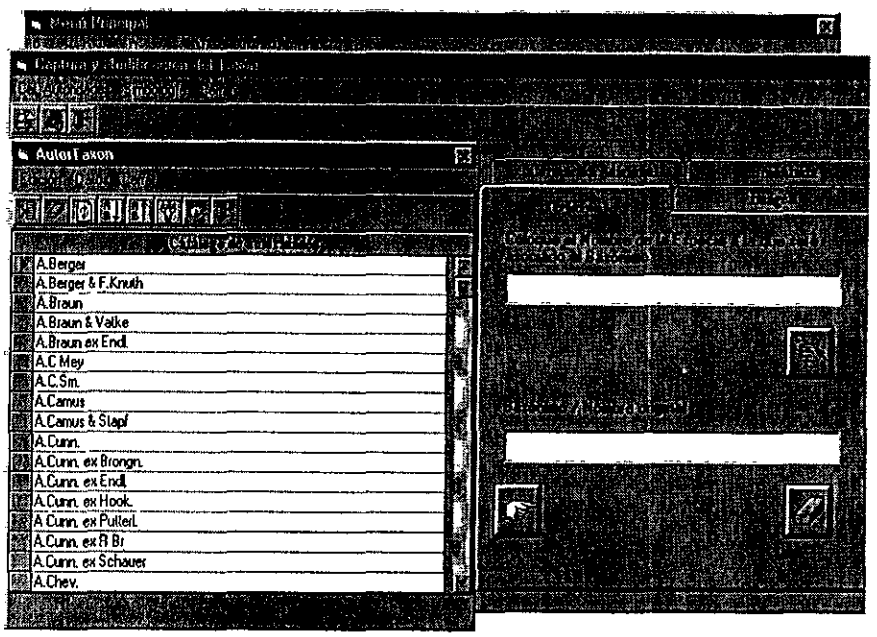


Figura 29. Pantalla del Catálogo de Autoridades.

Para dar de baja un taxón es necesario colocarse en la categoría taxonómica que se desea dar de baja, con un "*click*" del botón derecho del "*mouse*" aparece un menú de opciones donde se selecciona Baja del Taxón Señalado. Se eliminarán todos los nombres de las categorías taxonómicas inferiores asociadas jerarquicamente (figura 30).

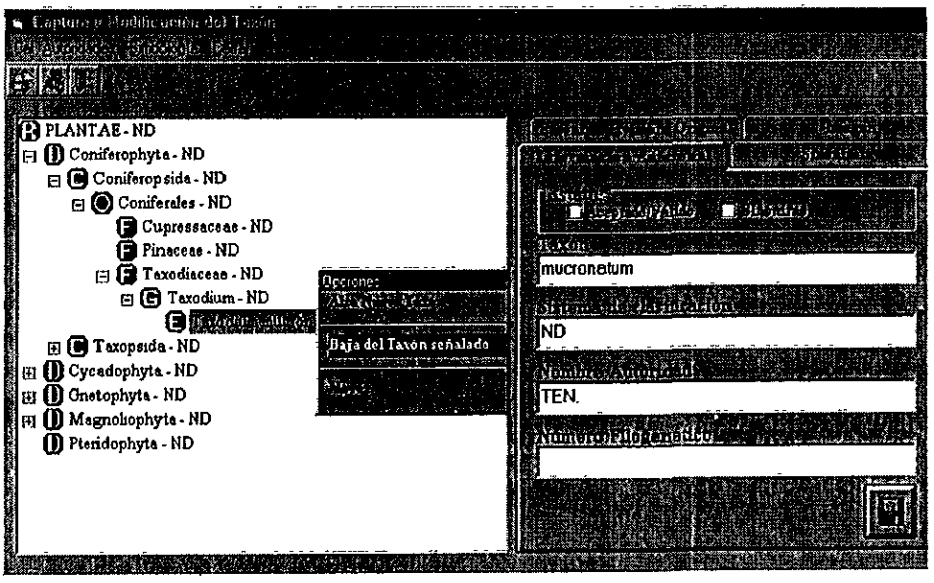


Figura 30. Pantalla Baja del Taxón Señalado.

Para mover un taxón de un nivel a otro, se debe colocar en la categoría taxonómica que se desea mover, habilitar el menú de opciones y elegir la opción Mover. Colocarse en el nivel inmediato superior donde se va a reubicar el nombre, y elegir la opción pegar del menú de opciones (figura 31).

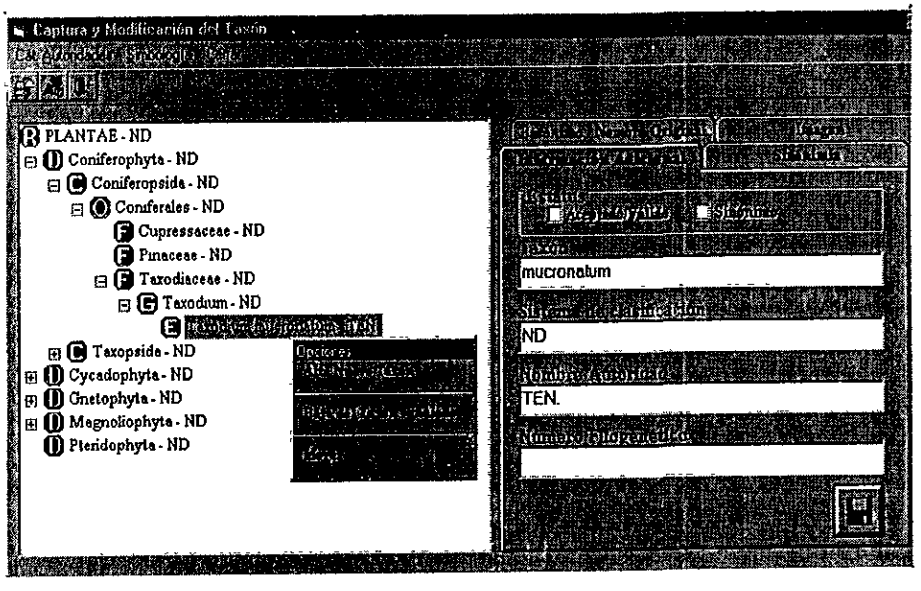


Figura 31. Pantalla para mover un taxón de un nivel a otro.

Para dar de alta una categoría taxonómica nueva utilice la pantalla Catálogo de Categorías Taxonómicas. Ubicarse en el nivel taxonómico próximo superior a la categoría que se va a dar de alta. Con el botón Nueva Categoría Taxonómica aparece un recuadro para teclear la categoría (figura 32).

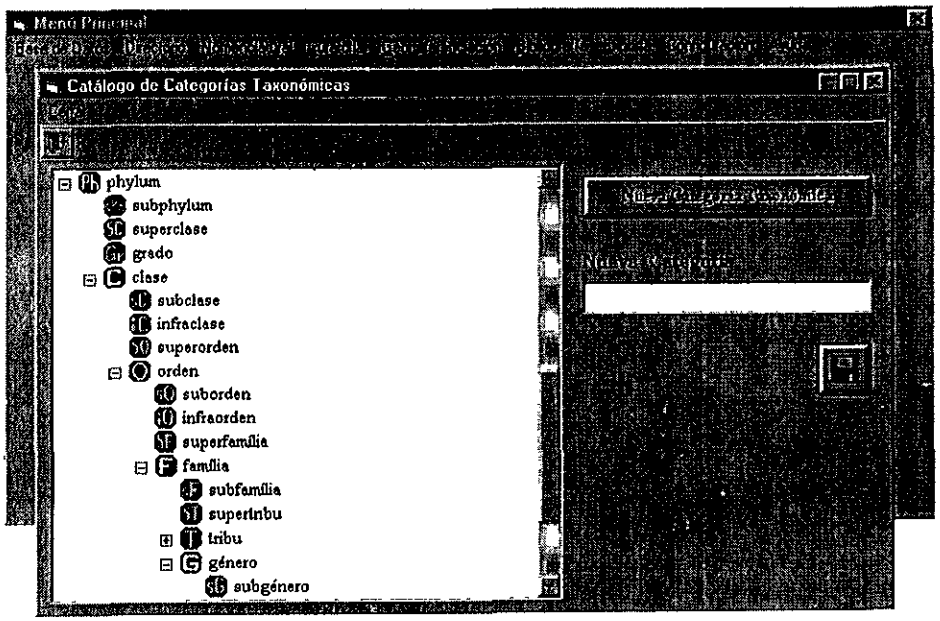


Figura 32. Pantalla del Catálogo de Categorías Taxonómicas para ingresar nuevas categorías.

Para dar de alta un nuevo taxón utilice la pantalla Captura y Modificación del Taxón. Colocarse en la categoría taxonómica próxima superior o igual al nivel del taxón que se va a dar de alta. Habilitar el menú de opciones y seleccionar Alta Nuevo Taxón (figura 33). En seguida se muestran los niveles taxonómicos que se pueden elegir.

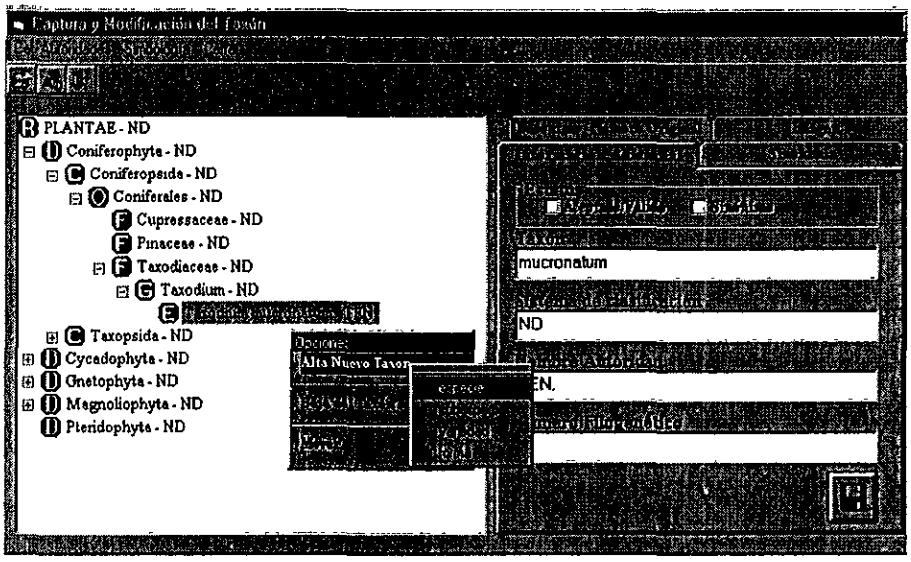


Figura 33. Pantalla de Alta Nuevo Taxón.

Entonces, se puede escribir el nombre que se va a dar de alta en la carpeta de Información adicional (figura 34).

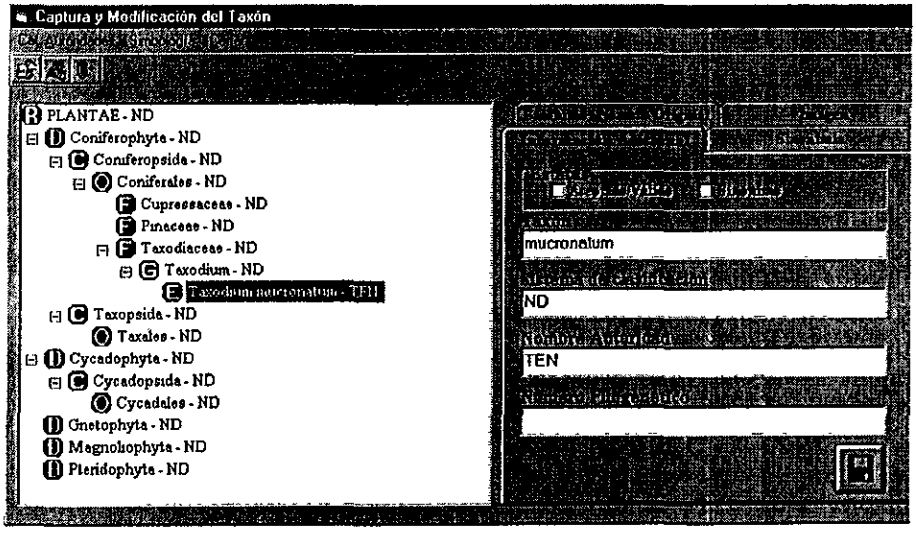


Figura 34. Pantalla de Captura y Modificación del Taxón para ingresar nuevos nombres de taxones.

El manejo de la información nomenclatural como se propone en este trabajo representa ciertas ventajas. Al analizar BIOTICA y compararla con otros modelo de bancos de datos de información curatorial se obtuvo lo siguiente:

5.4.2 BIOTICA y otros modelos de datos

- **PLATYPUS Versión 1.1** (Australian Biological Resources Study, Australian Nature Conservation Agency Canberra, Australia). Figura 35.

Este sistema para colecciones zoológicas integra los taxones superiores a familia a través de niveles definidos por el usuario (no preestablecidos): *Level 1* y *Level 2*. A partir del nivel *Family* las taxones que se pueden ingresar son: *Subfamily, Tribe Incertae sedis, Genus, Genus Available, Incertae sedis, Subgenus, Species, Species inquirenda, Species available,* y *Subspecies.*

Además, existe la posibilidad de integrar datos complementarios, por ejemplo, para el nivel de familia: introducción (texto en formato libre), última fecha de entrada de la información, captura de notas del catálogo, agradecimientos, nombres alternativos y detalles de compilaciones, ilustraciones, diagnosis, especies excluídas, y etiquetas opcionales.

- Ventajas de **PLATYPUS** sobre **BIOTICA**:

-Valida la entrada de información: no permite ingresar errores de captura como símbolos o terminaciones en las categorías.

-Una base de datos construída en Platypus puede ser exportada a un archivo de Access, *con algunas restricciones.*

-Elabora estadísticas de los elementos.

- Desventajas de **PLATYPUS** sobre **BIOTICA**:

-Está diseñado sólo es para clasificaciones zoológicas.

-El modelo de datos es similar al de BIOTICA, pero permite muchas formas de captura en el caso de los nombres: se pueden ingresar datos duplicados, un nombre binomial, etc.

-Contiene catálogos incorporados para las categorías taxonómicas, pero sólo para niveles por debajo de familia, para los superiores, no hay distinción (todos son *Level 1* o *Level 2*).

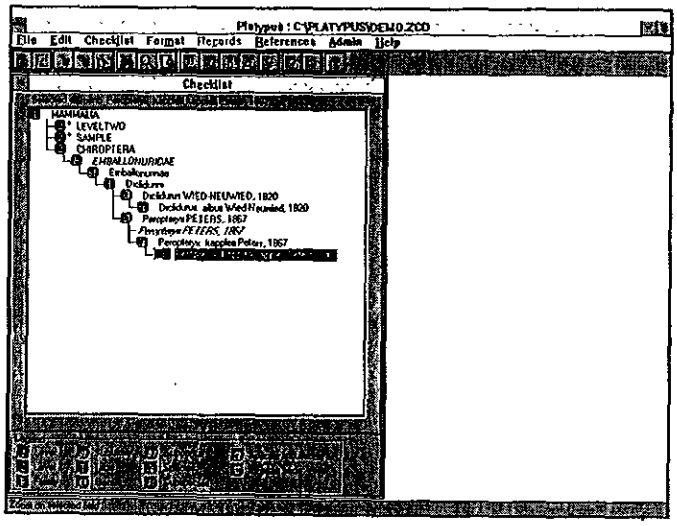
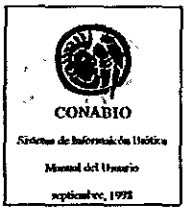


Figura 35. Pantalla de captura de **PLATYPUS Versión 1.1**.

BIOTA

A continuación se presenta un cuadro comparativo de la sección nomenclatural entre los modelos de BIOTA y BIOTICA (cuadro 13).

	BIOTA	BIOTICA
información Nomenclatural	Entidades: Species, Genus, Family Class, Phyllum, Kingdom	Entidades: Nombre, Categoría taxonómica Sinonimia, basónimos
	Autores a especie subespecie y variedad	Autores a cualquier nivel
	Captura de datos a cualquier nivel taxonómico (incluso sin relacionar)	*Se requiere reconstruir relación* en sistema de clasificación, *con el nombre ascendente* (no permite inconsistencias)
		

Cuadro 13. Comparación entre los modelos BIOTA y BIOTICA.

VI. CONCLUSIONES

1. La información nomenclatural asociada a colecciones biológicas científicas requiere un tratamiento especial al ser "modelada" en bases de datos debido principalmente a la complejidad de los datos y a la necesidad de representar fielmente una jerarquía taxonómica.

2. El empleo de catálogos controlados o diccionarios computarizados de taxones permite estandarizar y validar la información al someterla a un control de calidad.

3. La computarización de los sistemas de clasificación de los taxones más conocidos y estudiados en el país –plantas vasculares y vertebrados–, ha ayudado a eliminar la redundancia, duplicidad e inconsistencias en los datos, en la medida en que el catálogo provee la mayor parte de la información que requiere un taxónomo al identificar un organismo y darle un nombre. Este catálogo incluyó alrededor de 35,180 nombres de taxones de plantas vasculares y vertebrados de México, contenidos en diez sistemas de clasificación para más de 30 diferentes categorías taxonómicas, y cerca de 3,900 nombres de autores.

4. La idea de disponer de diversos sistemas de clasificación para un mismo taxón, un sistema eje y uno o más sistemas alternativos, ofrece la ventaja de validar más rápida y eficientemente la información ingresada a una base de datos. Este diccionario posee ocho sistemas de clasificación eje (para pteridofitas, angiospermas, gimnospermas, mamíferos marinos, mamíferos terrestres, anfibios, reptiles, aves y peces dulceacuícolas) y dos sistemas alternativos (para mamíferos terrestres y angiospermas).

5. El modelo entidad-relación permite modelar los elementos que integran un sistema de clasificación con sus atributos principales y las relaciones entre ellos. La combinación única de un nombre con sus ascendentes, la autoridad que lo publicó y el año de la publicación constituyen los atributos básicos del nombre de un taxón. Las relaciones incluyen vínculos jerárquicos entre un taxón en cierto nivel taxonómico con respecto a otros taxones en diferentes niveles taxonómicos, y vínculos entre autoridades y nombres de taxones.

6. La incorporación más eficiente de un sistema de clasificación a una base de datos es transladar un árbol inf-reticulado de la jerarquía taxonómica a una única entidad que contenga todos los atributos del nombre, haciéndose referencia siempre a ella misma y al mismo campo (relación unaria).

7. El modelo de datos de un diccionario nomenclatural computarizado debe permitir la recuperación del árbol completo del sistema de clasificación, independientemente del *software* y plataforma en el que esté construido.

VIII. LITERATURA CITADA

American Ornithologist's Union, 1983. *Check list of North American Birds by the Comittee on Clasification and Nomenclature*, USA.

Arnold, C., 1948. Classification of Gymnosperms from the viewpoint of Paleobotany. *Bot. Gaz.* 110 (1): 2-12.

Berendsohn, W., 1998. *A Workshop on the Compilation, Maintenance and Dissemination of Taxonomic Authority Files.* Washington D.C.

Bisby, F., 1993. *Botanical strategies for compiling a global plant checklist*, Chapter 14, *en* Bisby, F.A. *et. al.* Desings for a global plant species information system, The Systematics Association by Clarendon Press, Oxford, N.Y., Special Volume No. 48. p. 145-157.

Bold, H., C. Alexopoulos & T. Delevoryas, 1980. *Morphology of Plants and Fungi.* 4a ed. Harper & Row, Publishers, New York.

Bour, R. y A. Dubois, 1984. Nomenclature ordinale et familiale des tortues (Reptilia). Studia Geologica Salmanticoncia, Volumen Especial No. 1 (Studia Paleoqueliniologica 1): 77-86.

Brummitt, R.K. (*comp.*), 1993. *Vascular Plant Families and Genera.* Royal Botanic Gardens, Kew, Great Britain.

Ceballos, G. y H. Arita, 1996. Atlas de los mamíferos de México. Proyecto A003 de la CONABIO.

Código Internacional de Nomenclatura Zoológica, 1976. H. Blume Ediciones. Madrid.

CONABIO, 1998. *Comisión Nacional para el Conocimiento y Uso de la Biodiversidad*, México.

Corbet, G. & J. Hill, 1991. A *World List of Mammalian Species*, 3a ed., Oxford University Press. New York.

Cota, M., 1994. Ingeniería de Software, *Soluciones Avanzadas* No. 11, p. 15-16.

Crisci, J., A. Lanteri y E. Ortíz, 1994. Programas de computación en sistemática y biogeografía histórica: revisión crítica y criterios para sus selección. Cap. X. p. 207-225 *en* Llorente-Bousquets, J. e I. Luna (Comps.), 1994. *Taxonomía Biológica*, Ediciones Científicas Universitarias, México.

Cronquist, A., 1993. *A Commentary on the General System of Classification of Flowering Plants en* Flora of North America, North of Mexico, Vol. 1. Oxford University Press, New York, p. 272-293.

Dogra, P., 1980. *Embriologeny of Gymnosperms and Taxonomy*: An assessment *in*: P. K. K. Nair (*eds*). Glimpes in plant Research, 5, New Dehli, p. 114-128

Dubois, A., 1985. Miscellanea nomenclatorica batrachologica (VIII). Alyres, 4: 61-78.

Dundee, H., 1989. Higer category name usage for amphibians and reptiles. Systematic Zoology, 38: 398-406.

Engler, A., 1964. Syllabus der Pflanzenfailien. *Angiospermen* 12 :2 Alemania.

Espinosa, H., T. Gaspar, y P. Fuentes, 1993. *Listados Faunísticos de México III. Los Peces Dulceacuícolas Mexicanos.* Instituto de Biología, UNAM. México.

Flora of North America Editorial Committee (*eds.*), 1993. *Flora of North America* .Vol.1. Oxford University Press . New York.

Flores-Villela, O., 1993. *Herpetofauna Mexicana. Lista anotada de las especies de anfibios y reptiles de México, cambios taxonómicos recientes, y nuevas especies.* Carnegie Museum of Natural History, Pittsburgh.

Florin, C., 1955. *The Systematics of the gymnosperms en* A Century of progress in the Natural Sciences 1853-1953. Calif. Acad. Sci., San Francisco.

Flory, W., 1936. Chromosome numbers and Phylogeny in the Gymnosperms. *J. Arnold Arbor.* 17(2): 83-89.

Frost, D. R. (eds), 1985. Amphibian species of teh world, a taxonomic and geographical reference. Association of Systematics Collections, Lawrence, Kansas.

Gaussen, H. 1942-1955. *Les Gymnospermes actuelles et fossiles.* Fasc. 1-5: Géneralites. Trav. Lab. Forest. Toulouse, *I.* 2. Sect. 1, vol 1. Chap. 1-10 (Gymnosperms, extinct and extant orders, ill).

Green, D., 1994. Databasing diversity- a distributed, public domain approach, *Taxon* 43:51-62.

Hardin, J., 1971. Studies of the Southeastern United States flora. II. The gymnosperms. *J. Elisha Mitchell Sci. Soc.* 87: 43-50.

Hawks, C. A. and S. L. Williams, 1986. *Care of specimen labels in vertebrate research collections en* Proceedings of the 1985 Workshop on care and maintenance of Natural History Collections, *en* Ramírez-Pulido, J., I. Lira, S. Gaona, C. Müdespacher y A. Castro, 1986. Manejo y mantenimiento de colecciones mastozoológicas. UAM-I, México.

Hutchinson, J., 1924. Contributions towards a phylogenetic classification of flowering plants: III. The Genera of Gymnosperms. *Bull. Misc. Inform.* 2:49-66.

Jobling, J., 1991. *A Dictionary of Scientific Bird Names*, Oxford University Press.

King, F. W. y R. L. Burke, 1989. Cocodrilian, Tatuara, and Turtle Species of the World a Taxonomica and Geographic Reference. Association of Systematics Collections, Washington, D.C.

Koleff, P., 1997. Introducción a las bases de datos en la Biología Comparada Contemporánea, *Publicaciones Docentes del Museo de Zoología "Alfonso L. Herrera"*, Número 1, UNAM.

Llorente, J. y P. Koleff, 1997. La actividad taxonómica en México, *Biodiversitas*, 13:11-14.

Mabberley, D.J., 1987. *The Plant Book*. Cambridge University Press, USA.

Martínez, H., A. Castro y J. Ramírez, 1994. Avances en el uso de computadoras en las colecciones mastozoológicas. Departamento de Biología, UAM-Iztapalapa, Congreso Nacional de Mastozoología. *com. pers.*

Master Informática, 1992 *.dBASE III Plus*. Turgeon, Madrid.

McGee, M. and J. Boyce, 1993. *Microsoft Access for Windows, step by step.* Microsoft Access, USA.

Mickel & Beitel, 1988. Pteridophyte Flora of Oaxaca, México, *Mem. New York Bot. Gard.* 46: 1-568.

Microsoft Corporation, 1994a *Microsoft Access, Relational Database Management System for Windows, Version 2.0,* User's Guide, USA.

Microsoft Corporation, 1994b. *Microsoft Access 2.0 Introducción.* USA.

Microsoft Windows, página electrónica de Internet www.microsoft.com

Miller, A. H. *et al* (*eds.*), Cooper Ornithological Society, 1957. *Distributional Check-List of the Birds of Mexico. Part II.* Cooper Ornithological Society, California.

Minelli, A., 1993. *Biological Systematics.* Chapman & Hall, London.

Moran & Riba, 1985. *Flora Mesoamericana.* Vol 1:Pteridofitas

NABIN, 1997. *1997 Pilot Project.*

Nelson, J., 1993. *Fishes of the World,* 3a. ed., John Wiley & Sons, Inc:, New York.

Nielsen, E. & J. West, 1994. Biodiversity research and biological collections: transfer of information, Cap. 7 *en* P. Forey, C. Humphries, and R. Vane-Wright, 1994. Systematics and conservation evaluation, Systematics Association Special Volume No. 50, Clarendon Press, Oxford, p. 101-121.

Peláez, G. , 1994. *Taxonomía Biológica.* Cap. XII *en* Llorente-Bousquets, J. e I. Luna (*comps.*), Ediciones Científicas Universitarias, México.

Raddford, A; W.C.Dickison; J.R.Massey & R.Bell., 1974. *Vascular Plants Systematics.* New Yok, USA.

Ramirez, P., M.C. Britton, A. Perdomo y A. Castro, 1986. *Guía de los mamíferos de México,* Universidad Autónoma Metropolitana Unidad Iztapalapa.

Ramírez-Pulido, J., A. Castro-Campillo, J. Arroyo-Cabrales y F.A. Cervantes, 1996. *Lista Taxonómica de los Mamíferos Terrestres de México,* Museum Texas Tech University. 158:62.

Ramírez-Pulido, J., I. Lira, S. Gaona, C. Müdespacher y A. Castro, 1986. Manejo y mantenimiento de colecciones mastozoológicas. UAM-I, México.

Rob, P. and C. Coronel, 1995. *Database Systems. Design, Implementation, and Management.* 2nd Ed. International Thomson Publishing Company, U.S.A.

Salinas, M. y P. Ladrón de Guevara, 1993. Riqueza y Diversidad de los Mamíferos Marinos, *Ciencias*, 7:85-93.

Sivarajan, V.V., 1991. *Introduction to the Principles of Plant Taxonomy.* Second edition. Cambridge University Press, USA.

Soberón, J. and P. Koleff, 1998. The National Biodiversity Information System of Mexico. en Contribution to the Second National Forum on Biodiversity and Human Society: The Quest for a Sustainable Future. National Academic Press. *En prensa.*

Sporne, K., 1974. *The morfology of Gymnosperms. The structure and evolution of primitive seed plants.* 2nd ed. revised, Hutchinson Univ. Library, London.

Systematics Agenda 2000: Charting the Biosfera. Technical Report, 1994. Produced by Systematica Agenda 2000: a consortium of the American Society of Plant Taxonomists, the Society of Systematic Biologists at the Willi Hennig Society in cooperation with the Association of Systematic Collections, N.Y., USA.

Takhtajan, A.L. 1980. Outline of The Classification of Flowering Plants (Magnoliophyta). *The Botanical Review 46(3):226-267.*

Tryon & Tryon, 1982. *Ferns and Allied Plants,* Springer-Verlag

World Conservation Monitoring Centre (WCMC), 1997a. *Guide to the Darwin Initiative Handbooks on Biodiversity Information Management. Information and Policy.* Darwin Initiative Handbook No. 2, Cambridge, UK.

World Conservation Monitoring Centre (WCMC), 1997b. *Guide to the Darwin Initiative Handbooks on Biodiversity Information Management. Information Networks.* Darwin Initiative Handbook No. 4, Cambridge, UK.

World Conservation Monitoring Centre (WCMC), 1997c. *Guide to the Darwin Initiative Handbooks on Biodiversity Information Management. Data Custodianship and Access.* Darwin Initiative Handbook No. 7, Cambridge, UK.

Zhong. Y., S. Jung, S. Pramanik & J. Beaman, 1996. Modelo de datos y métodos de búsqueda y comparación para interaccion de clasificaciones en una base de datos taxonómica, *Taxon* 45: 223-241 .

Zug, G., 1993. Herpetology. *An Introductory Biology of Amphibians and Reptiles.* Academic Press, Inc. California.

Zellweger, C. & R. Allkin, 1993. *Alternative models for taxonomic data,* Chapter 25, *en* Bisby, F.A. et al Desings for a global plant species information system, The Systematics Association by Clarendon Press, Oxford, N.Y., Special Volume No. 48.

ANEXO 1

ETIQUETAS UTILIZADAS EN COLECCIONES CIENTÍFICAS

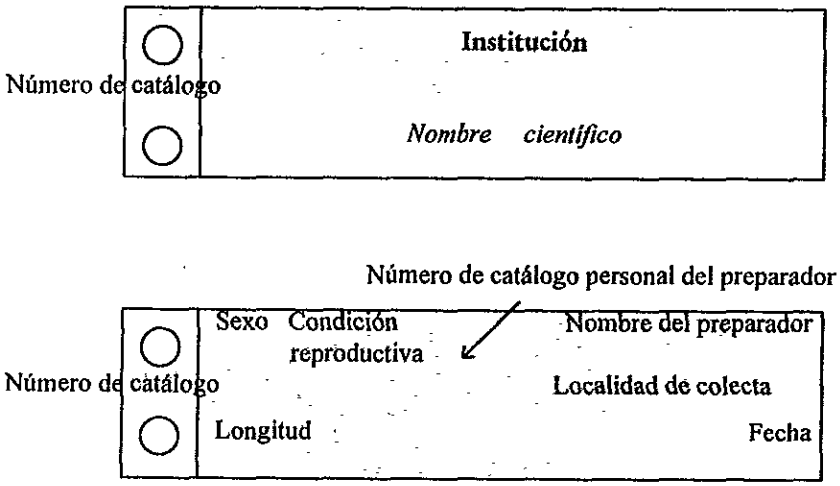


Figura 36. Ejemplo de etiqueta mastozoológica.

Museo de Zoología Facultad de Ciencias
UNAM
Los Chorros del Varal, Reyes de Salgado, Mich. 12-X-1990. A. Luis y L. Arteaga 900 msnm s/*Stevia*

Figura 37. Etiqueta para insectos preparados en alfiler.

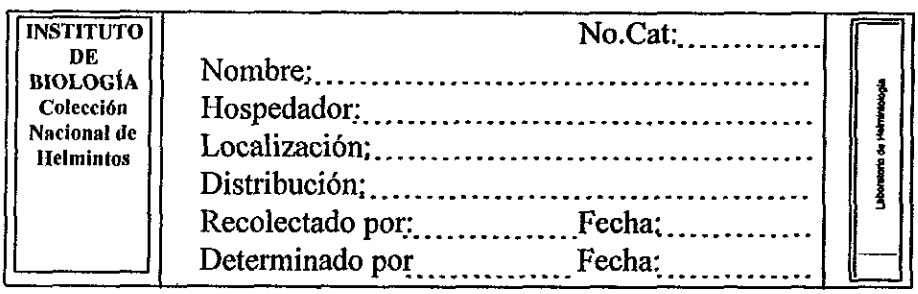


Figura 38. Etiqueta usada para helmintos parásitos en frasco.

Lab. Helmintología.

Núm:.....

C.N. HE

Figura 39. Etiqueta usada para helmintos en preparaciones microscópicas.

INSTITUTO DE BIOLOGIA U.N.A.M.
COLECCIÓN DE PECES

IBUNAM-P7785

Phycidae: 0
Alc 3, Esq 0, C & S 0

Urophycis cirratus

Atlantico norte
Tamaulpas

México
Golfo de México

Ident: DADZ/31-05-1995
Colect: L. Huidobro/E. Cruz
No colect: 10-89-53

25° 20' 1" N 96° 38' 9" W
09 Septiembre 1989
Medida: 209-218 mm

Figura 40. Etiqueta usada para peces.

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

MZFC Colección de Aves
Micrathene whitneyi

México: Querétaro. Peña Miller, 2Km E Peña Miller

Altitud: 1600 msnm 23-0ctubre- 1991

Peso:39g **Muda:** Abundante **Sexo:** H

Iris:amarillo **Osificación:** 100% **Hábitat:**Matorral xerófilo

Pico:verdoso **Grasa:**mediana

Patas: grises con filamentos **Preparó:** BEHB

Figura 41. Etiqueta usada para aves.

PLANTAS DE MÉXICO
HERBARIO NACIONAL DE MÉXICO (MEXU)
INSTITUTO DE BIOLOGÍA, UNAM

3689 Desmatodon cf. convolutus (Brid.) Grout.

Ladera basáltica con matorral xerófilo. Suelo cubriendo rocas, lugares expuestos y húmedos.

ESTADO DE MÉXICO. Municipio Tamamatla. 3 Km SW de San Pablo Atlazampam.
19° 21' N, 98° 56' W. elev. 2400 m.
Col. Angeles Cadenas Julio 7, 1984
Det. A. Cardenas Sep. 1992
Hojas con borde no diferenciado.

Figura 42. Etiqueta de herbario.

GLOSARIO

Aplicación de base de datos: una colección de herramientas las cuales facilitan el uso de una base de datos (WCMC, 1997a).

Atributo: propiedades de una entidad las cuales son mediciones para producir datos (WCMC, 1997a).

Atributos estandares: cierta información es esencial para todos los juegos de datos; otra información puede ser deseable pero no esencial (Green, 1994).

Autor: es la persona o personas quienes publicaron válidamente un nombre de un taxón (Zhong et al, 1996).

Autoridad: es la persona o personas quienes publican una clasificación que puede o no puede afectar el nombre usado para el taxón o el lugar taxonómico de un grupo (Zhong et al, 1996).

Base de datos relacional: una base de datos que consiste de dos o más tablas relacionadas vía campos comunes (WCMC, 1997a).

Base de datos de archivo plano (Flat-file): un tipo simple de base de datos que contiene solo una tabla de datos (WCMC, 1997a).

Base de datos: un conjunto de datos lógicamente estructurado y consistentes, generalmente manejados como archivos computarizados en uno o más sitios (WCMC, 1997a).

Campo: una columna en una tabla de una base de datos contiene valores de datos de una colección de datos existentes o interpretación de un proceso (WCMC, 1997a).

Calidad del dato: un estimado de la conveniencia de un juego de datos para ser usado en un propósito específico (WCMC, 1997a).

Clasificación: es la ordenación cientifica formal de las especies en un sistema jerárquico y la aplicación de nombres científicos a especies y grupos de especies (AS2000, 1994).

Clasificación: es un tratamiento taxonómico par un taxón particular o mas taxones con cierto rango. Categoria basada en datos de caracteres y asociados con un taxón relacionado (Zhong et al, 1996).

Diccionario de datos: Un deposito de información acerca de la definición y estructura de una base de datos (WCMC, 1997a).

Entidad: Un objeto de interés que consiste en una serie de atributos medibles (WCMC, 1997a).

Diagrama Entidad-Relación (E-R): Un diagrama que ilustra las interrelaciones entre una serie de entidades (WCMC, 1997a).

Dato: Es un hecho que resulta de observaciones o mediciones acerca del mundo, referido a alguna clase de estandar aceptado (WCMC, 1997a).

Estandares del dato: Métodos convenidos para el manejo del dato, los cuales reducen los costos de transacción y faculta la comparación de resultados en el espacio y el tiempo (WCMC, 1997a).

Hardware: El componente físico de uyn sistema de información tales como computadoras, impresoras, scáners, ploters y otros aparatos perifericos (WCMC, 1997a).

Información: Datos los cuales son analizados, integrados, o interpretados de otro modo para ser entendidos (WCMC, 1997a).

Lenguaje de consultas estructuradas (SQL): Lenguaje de consultas de base de datos extendido usado en muchos paquetes de bases de datos relacionales (WCMC, 1997a).

Nombre: la palabra o palabras que constituyen la designación científica de un taxón (Código Internacional de Nomenclatura Zoológica, 1976).

Nombre válido: el nombre correcto para un taxón dado; un taxón puede tener numerosos nombres utilizables, pero sólo uno entre ellos (lo más frecuente es que sea el más antiguo) es el nombre válido (Código Internacional de Nomenclatura Zoológica, 1976).

Normalización: Consiste en el proceso de lograr una estructura óptima para una base dedatos relacional (WCMC, 1997a).

Sistema de información: Conjunto organizado de personas, procesos, datos y herramientas para transformar los datos en información (WCMC, 1997a).

Querie: Una solicitud a una base de datos para selecionar y extraer datos (WCMC, 1997a).

Registro: Una colección de datos relacionados acerca de un tema, tratada como una unidad simple para propósitos de manejo o análisis de datos (WCMC, 1997a).

Relación: Se describe como dos o más entidades, estan relacionadas unas con otras (ej. ciertas especies pueden estar relacionadas a géneros por una relación de "pertenecer a" (WCMC, 1997a).

Servidor: Una computadora o programa que provee un servicio a otros programas o usuarios (WCMC, 1997a).

Sistema manejador de bases de datos relacional (RDBMS): Un sistema manejador de bases de datos el cual usa un modelo de datos relacional como su base (WCMC, 1997a).

Sistema manejador de bases de datos (DBMS): *Un paquete de software para manejo* de datos (WCMC, 1997a).

Sistemática: es la rama de la biología que busca el entendimiento de la diversidad de la vida (SA 2000, 1994).

Tabla: una entidad física en una base de datos relacional en la cual los datos se presentan en filas (registros) y columnas (campos) (WCMC, 1997c).

Taxón: toda unidad taxonómica que se emplee en biología (Código Internacional de Nomenclatura Zoológica, 1976).