



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

FACULTAD DE CIENCIAS

**SISTEMA DE CONSULTA DIGITAL PARA  
LAS COLECCIONES BIOLÓGICAS DEL  
INSTITUTO DE BIOLOGÍA DE LA U. N. A. M.**

**T E S I S**

QUE PARA OBTENER EL TÍTULO DE

**BIÓLOGO**

P R E S E N T A :

**OLIVER JOAQUÍN GIMÉNEZ HÉAU**



DIRECTOR DE TESIS: VÍCTOR HUGO REYNOSO ROSALES

2006



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

## Hoja de datos del Jurado

1. Datos del alumno  
Giménez  
Héau  
Oliver Joaquín  
53 35 04 35  
Universidad Nacional Autónoma de México  
Facultad de Ciencias  
Biología  
094556155
2. Datos del tutor  
Dr.  
Víctor Hugo  
Reynoso  
Rosales
3. Datos del sinodal 1  
Dra.  
Tila María  
Pérez  
Ortíz
4. Datos del sinodal 2  
Dra.  
Bertha Patricia  
Escalante  
Pliego
5. Datos del sinodal 3  
Dr.  
Fernando Alfredo  
Cervantes  
Reza
6. Datos del sinodal 4  
Dr.  
Enrique  
Martínez  
Meyer
7. Datos del trabajo escrito  
Sistema de consulta digital para las colecciones biológicas del Instituto de  
Biología de la U. N. A. M.  
63 p  
2006

**Agradecimientos:**

A mi asesor Víctor Hugo por su apoyo, tiempo y sobretodo confianza.

A mis sinodales: Tila María Pérez, Patricia Escalante, Fernando Cervantes y Enrique Martínez, por sus consejos y apoyo para llevar a cabo este trabajo.

A Karina por su amor, invaluable apoyo, dedicación y paciencia.

A Isabel Galina por su amistad y apoyo.

Al personal de la DGSCA por su apoyo técnico en la instalación de los portales.

A mis padres

A todos mis amigos

## Índice

1- INTRODUCCIÓN.....	3
1.1 - Las colecciones biológicas .....	5
1.2 - Historia de las colecciones biológicas de México.....	10
1.3 - Iniciativas internacionales para conectar colecciones biológicas .....	14
1.4 - Estándares internacionales .....	18
1.5 - Protocolos de comunicación.....	22
1.6 - Bases de datos .....	26
1.6.1 - Specify.....	27
1.6.2 - Biota .....	27
1.6.3 - Atta.....	28
1.6.4 - Biótica .....	29
1.7 - Biblioteca Digital .....	30
1.7.1 - Open Archives Initiative .....	30
1.7.2 - Dublín Core .....	32
2 - PROBLEMÁTICA.....	37
3 - OBJETIVO.....	37
4 - METODOLOGÍA .....	37
5 - RESULTADOS.....	38
5.1 - Extensiones al Darwin Core.....	38
5.2 - El portal DiGIR en el Instituto de Biología .....	40
5.3 - Una base de datos para el Instituto de Biología: recomendaciones .....	43
5.4 - OAI en el Instituto de Biología .....	47
6 - DISCUSIÓN .....	50
7 - ALCANCES A FUTURO.....	50
8 - CONCLUSIONES .....	53
ANEXOS .....	55
Anexo 1 - Campos definidos en el estándar internacional Darwin Core.....	55
Anexo 2 - Resultados de una consulta hecha a la Colección Nacional de Anfibios y Reptiles utilizando el Portal DiGIR.....	57
REFERENCIAS .....	60

## 1- INTRODUCCIÓN

El conocimiento se alberga en la memoria, sin embargo ésta no es suficiente para almacenar toda la información que conforma a un mundo tan extenso y complejo como el que nos rodea. Este es el motivo que dió origen a una de las más ambiciosas empresas del ser humano: la expansión de la memoria.

El desarrollo del lenguaje fue un primer paso en esta empresa, dando lugar así a la memoria colectiva alimentada por la transmisión oral del conocimiento. Probablemente casi todas las leyendas y relatos antiguos que se conocen en nuestros días tuvieron este inicio, como por ejemplo los poemas homéricos. Sin embargo, a pesar de que se utilizaban diversas artes nematotécnicas como la rima, para recordar y recitar los llamados libros orales, los relatos eran alterados por narradores, rapsodas, aedas o juglares, en función de sus intereses, distorsionando así el conocimiento original.

La escritura fue el siguiente paso para expandir la memoria humana, primero se copiaron objetos en las más diversas superficies: tablillas con cera, plomo, pieles, hueso, madera, papiros y amates. Pronto se comenzaron a representar símbolos, hasta que finalmente se llegó a la reproducción de sílabas.

En los primeros siglos d.C. apareció el codex, un predecesor del libro como lo conocemos hoy en día. Durante la edad media los libros eran copiados e ilustrados a mano por los monjes en los *scriptorium* (salas de escritura) de los monasterios, donde estaba prohibido hablar para evitar errores en las transcripciones, pero sobretodo estaba prohibido el uso de velas y cualquier otro tipo de fuego para evitar la pérdida de la información acumulada por un incendio.

A pesar de que en china se creó el primer libro impreso en el año 868 d.C. no fue hasta el siglo XV que los libros fueron accesibles para el resto de la población, con la aparición de la prensa de imprenta con partes removibles de Johannes Gutenberg en Alemania. Gracias a estos avances, durante los siguientes 400 años, los libros salieron de los monasterios e iglesias para instalarse en el mercado comercial, en donde fueron haciéndose cada vez más económicos y accesibles en todo el mundo gracias a nuevos inventos como el

correo, un paso más para lograr la expansión del conocimiento a un sector mas amplio de la población.

La ciencia es una de las áreas que se ha visto más beneficiada por esta expansión de la memoria en forma de publicaciones, y de hecho a sido una de sus grandes impulsoras buscando siempre nuevas formas de compartir eficientemente el conocimiento generado entre los científicos del mundo. Por ejemplo, la Segunda Guerra Mundial obligó a los científicos a trabajar muy estrechamente con el fin de conseguir armas cada vez más poderosas, creando así un precedente de colaboración científica que definiría nuevas formas de hacer ciencia, en donde las publicaciones académicas jugarían un papel fundamental.

Al terminar la guerra Vannevar Bush, asesor científico del presidente Roosevelt, preocupado por que la enorme cantidad de publicaciones científicas hacía imposible estar al día de todas las novedades, publicó un visionario artículo en 1945 llamado *As We May Think* (Bush 1945), en donde describe una máquina para gestionar grandes volúmenes de libros y documentos llamada Memex, abreviatura de *Memory Extended*. El memex tenía el tamaño de un escritorio y almacenaba los libros y otros materiales en microfilm. Tenía la capacidad de vincular documentos, almacenar las consultas realizadas y permitía que el usuario agregara notas. Desafortunadamente el Memex se quedó en una idea visionaria y nunca se desarrolló, pero influyó definitivamente en los nuevos conceptos que vinieron más adelante. Tal es el caso del hipertexto, acuñado por Ted Nelson en 1965, mismo que más adelante escribió un proyecto utópico llamado *Xanadu*, en el cual todos los textos del mundo se almacenarían permanentemente en un depósito universalmente accesible, una especie de memoria o acervo colectivo universal. Tres años después otro visionario, Alan Kay, creó el modelo del Dynabook, la primer computadora portátil. Kay llamó acertadamente al dispositivo un "amplificador de la mente".

Unos años antes apareció por primera vez el término "base de datos", en un simposio que llevaba el nombre de *Development and Management of Computer-centered Data Base*. Se dice que las bases de datos tuvieron su origen en el proyecto Apolo, ya que era imposible para la mente humana organizar todas las variables que debían de tomarse en cuenta para mandar a un hombre a la Luna.

El mismo año en que el astronauta dejaba su huella en la Luna, el Pentágono conectaba el primer nodo del ARPANET, la primera red de computadoras del mundo. Gradualmente ARPANET fue conectándose con otras redes de computadoras, pero fue hasta 1983 cuando se considera que nació realmente la Internet, cuando se separó la parte militar de la civil de la red. En ese momento ya la componían 500 servidores.

Poco antes había aparecido la primera biblioteca digital en línea, se llamó “Proyecto Gutenberg” era de acceso libre y llegó a reunir 2000 ejemplares digitalizados de la literatura universal.

A partir de este momento, la comunidad científica adoptó rápidamente las bases de datos, las publicaciones electrónicas y el Internet para almacenar y transmitir el conocimiento que generaban, culminando hoy en día en complejos sistemas de publicación digital. De esta manera los seres humanos no solo hemos logrado expandir la memoria, sino también la mente, para poder así comprender mejor los fenómenos y objetos que nos rodean en la naturaleza.

Las bibliotecas han acompañado a la escritura a lo largo de su historia, desde las bibliotecas que almacenaban ladrillos grabados con escritura cuneiforme de la cultura asiria, hasta las bibliotecas virtuales que hoy en día proliferan, como Google Books, que almacena millones de libros en formato digital, accesibles desde cualquier parte del mundo con una computadora y una conexión a Internet.

Las bibliotecas son colecciones de libros o documentos, y funcionan con los mismos principios que muchas otras colecciones, como las que se analizarán en este trabajo. Las colecciones biológicas, son más específicas que las bibliotecas, pero con muchos rasgos comunes. Durante siglos, muchos naturalistas y científicos han depositado en éstas millones de muestras preservadas de la biodiversidad de nuestro planeta. Hoy en día las colecciones biológicas enfrentan un nuevo reto: integrar los nuevos sistemas de información digital para convertirse en herramientas aún más poderosas.

### **1.1 - Las colecciones biológicas**

Se calcula que las colecciones biológicas del mundo almacenan entre 2 y 3 mil millones de ejemplares biológicos distribuidos en todo el planeta (Krishtalka y Humphrey 2000). La suma de estas colecciones es la mejor documentación

que existe de la composición global, la identidad, la distribución espacial, la ecología, la sistemática y la historia de las formas de vida conocidas en el planeta.

Los ejemplares de las colecciones biológicas han sido recolectados en todo el mundo durante cientos de años, ofreciendo una visión única de la biodiversidad a lo largo del tiempo, incluso de la que existía hace millones de años, gracias a la existencia de colecciones paleontológicas que almacenan organismos fósiles. Las colecciones biológicas proveen información sobre la presencia y ausencia de especies a lo largo del tiempo y en el espacio, además de información geográfica, climática, geológica y algunos otros datos de las regiones donde fueron colectados (Peterson y Soberón 2004). Cada ejemplar está acompañado por información que permite ubicarlos en un lugar del planeta (sitio de colecta), de la biodiversidad (información taxonómica) y del tiempo (fecha de colecta) (Winker 2004). De algunos ejemplares, principalmente mamíferos y aves, inclusive puede obtenerse información de sus hábitos alimenticios, al analizar sus contenidos estomacales o de algunas ciertas enfermedades que dejan vestigios en su piel o plumaje (Winker 2004).

Los ejemplares contienen además mucha información que no se conoce en el momento de su colecta; sin embargo, con el avance en las tecnologías y conocimiento se pueden extraer estos datos posteriormente. Tal es el caso de la información geonómica o molecular, que puede ser extraída hoy en día de ejemplares colectados hace más de cien años cuando no se conocían estos conceptos. Probablemente en el futuro se encontrará información que se desconoce hoy en día en los ejemplares de las colecciones (Suarez y Tsutsui 2004), por este motivo es fundamental su correcta conservación.

Durante los últimos doscientos años, las colecciones han sido fundamentales para el desarrollo de la biología. El auge de las colecciones biológicas proviene de las épocas de las grandes exploraciones del planeta, su primer función fue la de dibujar la taxonomía de los grupos biológicos que habitaban las regiones exploradas. Esta misión se llevó a cabo relativamente bien. Pero a partir del momento en que es menester de la ciencia monitorear los cambios que se presentan en la biodiversidad, como consecuencia de las acciones del hombre, las colecciones adquieren nuevamente un papel principal que no debe ser descuidado (Winker 2004).

Muchas veces las colecciones son comparadas con bibliotecas de la vida (Suarez y Tsutsui 2004), haciendo una analogía entre los ejemplares preservados de las primeras con los libros de las segundas. A diferencia de las bibliotecas donde existen muchos ejemplares idénticos, en las colecciones biológicas cada ejemplar es único, además de que en las colecciones los ejemplares cambian de catalogación según se modifique la taxonomía; no hay que olvidar que se trata de un área de la ciencia que está en constante cambio.

Muchos ejemplares son comúnmente usados en investigaciones, y deben ser catalogados y almacenados cuidadosamente con el objetivo de que en el futuro puedan ser revisados para confirmar su identificación por otros investigadores. Muchos artículos científicos publicados hoy en día basan sus datos en información obtenida de los ejemplares de las colecciones biológicas, en particular en las áreas de sistemática, filogeografía, biogeografía, ecología y biodiversidad. De hecho el desarrollo y crecimiento de las colecciones biológicas en los últimos años ha sido fuertemente incentivada por la publicación de artículos en estas áreas.

La vida útil de los ejemplares de las colecciones es mucho más larga que la vida útil de un artículo científico escrito sobre ellos. Solamente un tercio de las publicaciones científicas exceden los 10 años de vida media (Winker 2004), mientras que los ejemplares de las colecciones pueden durar cientos de años y seguir brindando información relevante, en la misma u otra disciplina. De esta manera hoy en día las colecciones biológicas se utilizan no sólo como un acervo de información biológica, sino también como una herramienta de la investigación científica, contribuyendo de manera única e invaluable en el estudio de los patógenos, vectores de enfermedades y contaminación ambiental, así como para medir la biodiversidad y su pérdida, invasiones biológicas y cambio climático global (Suarez y Tsutsui 2004).

Las colecciones han sido usadas comúnmente para descubrir la historia de enfermedades infecciosas e identificar sus fuentes y reservorios. Este es el caso de las pieles de simios africanos (*Cercocebus torquatus*), almacenadas en el Smithsonian Institution desde 1800 que han sido usadas recientemente para rastrear al SIVsm, un síndrome de inmunodeficiencia de primates muy similar al producido por el VIH, encontrando su presencia desde 1896. Estos estudios abren nuevas posibles teorías acerca de la aparición de esta enfermedad en

humanos. También se han utilizado millones de ejemplares de mosquitos almacenados en todo el mundo para obtener información de las dinámicas de poblaciones y patrones de transmisión de muchas enfermedades portadas por éstos, como la Malaria Aviar o el Virus del Nilo entre muchas otras (Fonseca et al. 2001; Suarez y Tsutsui 2004).

A finales de los años 90 el virus de la influenza hallado en tejidos preservados de humanos fallecidos durante la pandemia de esta enfermedad en 1918, fue comparado con el virus de la influenza aviar, encontrado en ejemplares de esa misma época, colectados y almacenados en las colecciones biológicas de los museos. Gracias a estos estudios se comprobó que la enfermedad no había sido transmitida por las aves, como se pensaba, sino por los puercos (Taubenberger, Reid et al. 1997). Otros estudios han reconstruido la historia evolutiva de los virus, proveyendo una gran guía para el desarrollo de vacunas eficaces (Suarez y Tsutsui 2004).

Un caso más en el que las colecciones han sido claves para entender la dinámica de las enfermedades humanas es el caso del virus del Hanta. Estudiando ejemplares preservados de roedores del suroeste de los Estados Unidos, este virus se asoció a la explosión de las poblaciones de *Peromyscus maniculatus*, un roedor silvestre, después del fenómeno del Niño en 1992 con la presencia de una misteriosa pulmonía que acabó con la vida del 70% de los humanos infectados. (Yates, Mills et al. 2002).

Respecto al monitoreo ambiental, las colecciones biológicas se pueden usar para medir las cantidades históricas de contaminantes presentes en el pasado y poder compararlas con las actuales. Por ejemplo, se han utilizado ejemplares preservados de aves marinas de algunos países como Suecia para demostrar que fue por causa de la industrialización humana de los años cuarenta y cincuenta que aumentaron enormemente las concentraciones de mercurio en la atmósfera, afectando las cadenas alimenticias (Monteiro y Furness 1997; Berg, Johnels et al. 1966).

Las colecciones han sido utilizadas constantemente para detectar el impacto que ocasiona el ser humano sobre otras especies, por ejemplo, el estudio de huevos preservados en colecciones permitió descubrir la relación que existe entre el uso indiscriminado de DDT y la reducción de poblaciones de muchas especies de aves. Los hallazgos realizados en estos estudios

impulsaron importantes cambios en las legislaciones que beneficiaron a los ecosistemas (Winker 2004). Se ha demostrado también que las deformaciones de las ranas aumentaron en gran medida después del uso masivo de los herbicidas (Suarez y Tsutsui 2004). Finalmente, los ejemplares de colecciones, han permitido descubrir un gran aumento en las tasa de mutación de las golondrinas después de la explosión en 1986 de la planta nuclear de Chernobyl (Ellegren, Lindgren et al. 1997).

En las colecciones se encuentra el registro del cambio de las poblaciones debido a la pérdida de hábitat, a la fragmentación, a los invasores biológicos y al cambio climático global. Por ejemplo, en Kenya se han usado aves de las colecciones biológicas para demostrar que los rasgos fenotípicos de las aves pueden ser buenos para predecir la persistencia de una especie en zonas ecológicamente fragmentadas (Lens, Van Dongen et al. 2002). Utilizando ejemplares de colecciones, también se ha estudiado cómo afectaría el cambio climático global a la distribución y posible extinción de distintas especies, (Peterson, Martinez-Meyer et al. 2004 ; Peterson, Ortega-Huerta et al. 2002 ; Thomas, Cameron et al. 2004; Martinez-Meyer, Peterson et al. 2004; Peterson y Siqueira 2003).

La información contenida en las colecciones biológicas también puede tener importantes repercusiones económicas. Por ejemplo, en los Estados Unidos, las plagas de artrópodos cuestan a los graneros arriba de 14 mil millones de dólares (Suarez y Tsutsui 2004). Las colecciones se pueden utilizar para predecir de dónde y cuándo aparecerán plagas (Sanchez-Cordero y Martinez-Meyer 2000; Soberon, Golubov et al. 2001; Costa, Peterson et al. 2002).

Por otro lado los ejemplares de las colecciones han sido usados para predecir pestes y patógenos o vectores de enfermedades (Costa, Peterson et al. 2002; Peterson, Sanchez-Cordero et al. 2002; Peterson y Shaw 2003) como por ejemplo la modelación de la distribución de los roedores hospederos y de las chinches *Triatoma*, vectores del *Trypanosoma cruzi*, protozooario responsable de la enfermedad de Chagas, en México, con el fin de poder ubicar las zonas donde el riesgo de contraer la enfermedad es mayor. Estos trabajos pueden traducirse en importantes ahorros para el sistema de salud, ya que se pueden enfocar los recursos económicos a las zonas donde los

organismos que forman el ciclo de vida de la enfermedad están presentes simultáneamente.

Finalmente, las colecciones biológicas han permitido también predecir la distribución de especies conocidas y desconocidas. Por ejemplo, en Madagascar se encontró una nueva especie de camaleón al modelar el nicho ecológico a partir de ejemplares de colecciones de una especie hermana (Raxworthy, Martinez-Meyer et al. 2003). De manera similar se han realizado estudios en la predicción de especies invasoras (Peterson y Vieglais 2001).

En los últimos años ha quedado demostrada la gran importancia de las colecciones biológicas en áreas emergentes de la ciencia, dándole un nuevo valor a los ejemplares colectados en el pasado. Por este motivo los países que cuentan con un rico acervo de biodiversidad en sus distintas colecciones tienen la oportunidad de utilizarlo para investigaciones de frontera y desarrollar nuevo conocimiento a partir del trabajo realizado durante décadas por sus taxónomos.

## **1.2 - Historia de las colecciones biológicas de México**

México cuenta con una gran tradición naturalista, comenzando por la admiración que tenían las culturas prehispánicas a la naturaleza, creando jardines botánicos con especies de todo el continente, por la muy desarrollada cultura herbolaria de los grupos indígenas y por los grandes viajes que realizaron en México naturalistas de todo el mundo.

La primera colección de plantas mexicanas de la que se tiene registro, fue creada en 1729 por un holandés llamado William Houstoun. Sin embargo, hoy no permanece ningún ejemplar en México ya que ésta fue repartida entre varios herbarios del mundo principalmente al jardín de Chelsea, en Inglaterra. Posteriormente se tienen registros de las colecciones realizadas por Luis Née, nacionalizado español de origen francés, quien llegó a Acapulco después de un viaje a Centro América, motivado por el apoyo que brindó Carlos III a la ciencia. Née logró coleccionar 10,000 ejemplares en su viaje del puerto a la Ciudad de México. Desafortunadamente éstos ejemplares se localizan hoy en día en Praga, Viena y Kiev.

Un contemporáneo de Née, Martín Sessé, convenció al Rey de España para organizar una expedición botánica en México. El Rey accedió y lo nombró director de dicha expedición con la orden de formar un jardín botánico. Así

nació la *Expedición Botánica de la Nueva España*, la cual realizó ocho viajes por todo el país entre los años de 1788 y 1795. Con colectas de este viaje, describieron 2500 especies de plantas y realizaron 2000 ilustraciones botánicas en la Ciudad de México. Las obras fueron enviadas a Madrid, pero su portador, Mariano Mociño, prefirió depositarlas en Montpellier debido a la inestabilidad política en la que se encontraba España en esa época. Las plantas fueron entregadas a De Candolle, quien posteriormente fundó el Jardín Botánico de Ginebra, en el cuál dedicó una sección a plantas mexicanas. Desafortunadamente las colectas se perdieron durante muchos años, pero hoy en día se sabe que se encuentran en el Hunt Institute for Botanical Documentation, en los Estados Unidos de Norteamérica.

En la expedición se colectaron también plantas y animales de todo el país, la mayoría de estos ejemplares se depositaron en el Herbario y en el Museo de Historia Natural de Madrid y otros fueron puestos bajo la custodia de Vicente Cervantes, profesor de la cátedra de botánica en la Ciudad de México, quien los guardó en un herbario que existió de 1791 a 1820 en una esquina de Palacio Nacional. Hoy en día, en el Herbario Nacional, se conservan 63 ejemplares de esta colección, compuestos por líquenes y fanerógamas. El resto de los ejemplares colectados durante la expedición botánica se encuentran actualmente en el Jardín Botánico de Madrid, en el Field Museum de Chicago, en el Fielding Druce Herbarium de Oxford y algunos cuantos en la Universidad de Puebla, en México.

El Jardín Botánico encomendado por el rey a Sessé, fue inaugurado el 20 de diciembre de 1788, junto con la cátedra de botánica que se incorporó a la carrera de medicina en la Real y Pontificia Universidad de México. Con el desarrollo del curso los alumnos fueron colectando plantas, dando lugar al primer herbario mexicano con 200 ejemplares en 1789. Ese mismo año llegaron a México las primeras obras botánicas impresas, algunas de las cuales aún se conservan en el Acervo Histórico del Herbario Nacional.

El hijo de Vicente Cervantes, Julián Cervantes, además de publicar la primer obra botánica mexicana (*Tablas Botánicas que, para el más pronto y fácil estudio de esta ciencia dispuso el presbítero D. Julián Cervantes*), dejó un herbario de 1700 ejemplares que aún se conserva en la Universidad de Puebla, aunque diez ejemplares están resguardados en el Herbario Nacional.

Durante el periodo presidencial de Guadalupe Victoria, en 1825, todas las colecciones se reunieron en el Museo Nacional de la Universidad, junto con las del Conservatorio de Antigüedades y de Historia Natural, así como con la Biblioteca Nacional. Dos años más tarde, se fundó ahí el Herbario Nacional.

En 1849 se organizó la Comisión de Límites entre México y Estados Unidos, en la que se encontraba el botánico Charles Parry quien recolectó plantas en Texas, Tamaulipas y Nuevo León, éstas se encuentran hoy en día en el acervo general del Herbario Nacional.

Otra comisión de gran importancia dentro de la historia de las colecciones nacionales, fue la Comisión Científica del Valle de México, en la que se integró un herbario, una colección de maderas y una de aves. Estas colectas son de gran interés en nuestros tiempos ya que provienen de sitios que hoy se encuentran urbanizados como Iztapalapa, el Ajusco, Mineral del Chico, Real del Monte y Xochimilco.

Quince años después, en 1865 y por decreto de Maximiliano, se fundó el Museo Nacional de Historia Natural, Arqueología e Historia. El interés que presentaba Maximiliano para fomentar la cultura, llevó éste museo a Palacio Nacional, donde quedó bajo el cuidado de su jardinero personal Dominique Bilimek, quien incorporó muchos ejemplares botánicos provenientes de Europa.

Para 1877, año en el que el museo contaba con 90,000 ejemplares de plantas, animales y rocas, se formó la Comisión Geográfica Exploradora, ésta creó un mapa geológico, distintos listados faunísticos nacionales y varias colecciones de flora y fauna. Estas colecciones pasaron a formar parte del acervo del recién creado Museo de Historia Natural, sucesor del Museo Nacional de Historia Natural, Arqueología e Historia.

Finalmente, en 1882 se formó la Comisión Científica Mexicana, dirigida por Alfonso L. Herrera, con el objetivo de explorar y coleccionar ejemplares tanto botánicos como zoológicos de toda la república. Herrera trabajó en el Instituto Médico Nacional, la institución de investigación más importante de su época, en donde se realizaban estudios principalmente con las plantas medicinales, de las cuales se obtenían extractos que eran probados en animales y humanos. De esta colección se conservan muestras de 50 especies dentro del Acervo Histórico del Herbario Nacional. El Instituto Médico Nacional existió hasta 1915, albergando una colección de 16,000 ejemplares, que fueron trasladados al

recién creado Instituto de Biología General y Médico, dirigido por Alfonso L. Herrera. Herrera organizó excursiones semanales con el fin de incrementar el acervo de las colecciones, y en 1922 ingresó ejemplares de importancia económica, además de los de importancia medicinal.

Este nuevo Instituto reunió a las colecciones del Museo Nacional de Historia Natural, del Instituto Médico Nacional y de la comisión Geográfica Exploradora, además de administrar un jardín botánico, un parque zoológico, una estación de biología, y un acuario de agua dulce, también se pensó en construir un acuario de agua salada, pero nunca se realizó. Junto con el Museo de Historia Natural y el Departamento de Exploración de la Flora y Fauna, el Instituto forma la Dirección de Estudios Biológicos.

Con la creación de la Universidad Nacional Autónoma de México, la Dirección de Estudios Biológicos se convirtió en el Instituto de Biología, con sede en la Casa del Lago, en Chapultepec, incluyendo al Museo de Historia Natural, localizado en la calle de Chopo. Las colecciones biológicas fueron muy beneficiadas en esta nueva sede, por ejemplo el Herbario Nacional que en 1937 contaba con 28,000 ejemplares en 1953 contó con 70,000.

En 1941 se integró al herbario el botánico español Faustino Miranda, quien incorporó su colección de 10,000 ejemplares colectados en Puebla, Guerrero, Morelos, Chiapas y Yucatán. De la misma manera, en 1951 se incorporó Eizi Matuda, botánico japonés, quien reunió una colección de aproximadamente 28,000 ejemplares.

En 1980 se inició un plan de recolecta de la flora mexicana ampliando el herbario con ejemplares de zonas poco representadas. Mario Sousa incorporó también importantes colecciones particulares como la de los hermanos lasallistas o como la de Wolfgang Boege, quien había realizado importantes colectas en el estado de Puebla (Dávila Aranda y Germán Ramírez 1991). También es Sousa quien incorporó las primeras computadoras para la elaboración de etiquetas en el Herbario.

Las colecciones zoológicas también se vieron incrementadas por campañas de colecta o donaciones de colecciones particulares, como sucedió con la colección Anita Hoffmann, una muestra muy representativa de la acarofauna mexicana, y la integración de importantes investigadores como el Dr. Enrique Rioja Lo Bianco en la colección carcinológica y el Prof. Carlos C.

Hoffmann en la colección entomológica (Brailovsky y Gómez Varela 1993). En 1979, las colecciones adquirieron el estatus de institucionales y se nombraron curadores responsables de cada una de ellas. Hoy en día el Instituto de Biología cuenta con diez Colecciones Nacionales zoológicas (Helmitológica, Malacológica, Acarológica, Carcinológica, Entomológica, Ictiológica, Herpetológica, Ornitológica y Mastozoológica) y el Herbario Nacional.

México, como país megadiverso, se encuentra en los primeros lugares de biodiversidad de anfibios, reptiles, mamíferos y plantas. Esta gran diversidad se encuentra relativamente bien documentada en las colecciones biológicas de instituciones nacionales e internacionales, y en particular en las colecciones biológicas del Instituto de Biología de la UNAM. Dada la importancia que puede tener la información contenida en dichas colecciones, tanto para la ciencia básica (e.g. sistemática, biogeografía) y aplicada (detección de plagas y enfermedades) como para la conservación de la diversidad de nuestro país, es crucial eficientizar el manejo de la información contenida en las colecciones biológicas.

### **1.3 - Iniciativas internacionales para conectar colecciones biológicas**

Las colecciones biológicas son una herramienta importantísima para la investigación científica, no sólo como un catálogo de la vida en los distintos continentes, sino como un acervo de información pasada y presente sobre la biodiversidad del planeta, el clima y el ambiente. En ellas se almacena una gran cantidad de información fundamental para la conservación del planeta. Por este motivo, es fundamental no sólo conservarlas, sino integrarlas a los nuevos sistemas tecnológicos de información, como el Internet, con el fin de que cada vez más científicos de todo el mundo puedan acceder a estos datos, lo que aumenta la posibilidad de encontrar nueva información relevante para la conservación y la salud.

El reto de conservación de la biodiversidad al que se enfrenta la ciencia en nuestros tiempos requiere, además de información, tecnología y profesionales de mucha cooperación internacional para lograr entender el complejo tejido que es la vida en nuestro planeta. Afortunadamente en las últimas décadas se han realizado esfuerzos importantes por lograr esta cooperación, destacando la Cumbre de Río, llevada a cabo en 1992.

En la Cumbre de Río se alcanzaron dos acuerdos de suma importancia, ambos muy relacionados entre si: la Convención de las Naciones Unidas para el Cambio Climático Global y la Convención de la Diversidad Biológica (CDB). La primera buscaba conseguir la estabilización de los gases causantes del efecto invernadero, mientras que la segunda buscaba la conservación y el uso sustentable de la diversidad biológica. Ambas requirieron de la colaboración de muchas naciones y organizaciones tanto nacionales como internacionales, integrando toda la información que se encuentra distribuida en éstas, además del desarrollo de áreas emergentes en la ciencia como lo es la informática de la biodiversidad, que combina el desarrollo de las tecnologías de información y bases de datos con la información biológica almacenada durante cientos de años en los museos de historia natural.

Destaca el artículo 17 de la CDB donde se menciona la necesidad del intercambio de toda la información de las fuentes públicas disponibles, relevantes para la conservación y uso sustentable de la biodiversidad, incluyendo reportes técnicos, investigaciones científicas y socioeconómicas.

Para apoyar el acceso a la información, desde el nivel de genes hasta los ecosistemas y desde los reportes técnicos hasta las investigaciones científicas y socioeconómicas que se encuentran distribuidas por todo el mundo, la CDB implementó la Clearing House Mechanism (CHM), una red basada en Internet que promueve la cooperación técnica y científica en el intercambio de la información.

Australia fue el primer país en desarrollar sistemas para el intercambio de información sobre la biodiversidad. Desde mediados de los años setentas, los herbarios australianos comenzaron a trabajar en proyectos para digitalizar sus datos de una forma cooperativa. The Environmental Resources Information Network (ERIN) se estableció en 1989 para proveer información ambiental georreferenciada, con la que se podía planear y apoyar la toma de decisiones, ese mismo año se desarrolló también el Protocolo y Estándar de Información para el Intercambio de Datos de los Herbarios. Un estándar diseñado para el intercambio electrónico de información sobre biodiversidad. ERIN fue un ejemplo que siguieron la CONABIO en México, la INBio en Costa Rica y el proyecto brasileño Base de Datos Tropical.

A principios de los noventa, científicos de diversas áreas crearon el Biodiversity Information Network – Agenda 21 (BIN21) que se configuró como una red informal de colaboración consistente en la participación de una serie de nodos de todo el mundo. Por otro lado, en 1998 se inició en el museo de historia natural de la Universidad de Kansas un proyecto llamado Species Analyst cuyos objetivos eran desarrollar estándares y herramientas de cómputo para acceder y analizar el mundo de las bases de datos de colecciones biológicas y de observaciones. Este proyecto fue uno de los primeros en el mundo en usar sistemas de bases de datos distribuidas de colecciones biológicas en el mundo. A partir de este momento se continuó el desarrollo del GARP (Generic Algorithm for Rule-set Prediction), una herramienta de cómputo usada para predecir la distribución de especies a partir de los ejemplares conservados en las colecciones. El GARP fue creado originalmente por David Stockwell, en la Universidad de San Diego (Stockwell and Nobel 1991).

Uno de los esfuerzos más notables para facilitar el acceso a la información de la biodiversidad en el mundo es el Global Biodiversity Information Facility (GBIF). Surgió a partir de la recomendación del grupo de trabajo del Foro de Megaciencia de la Organización para la Cooperación Económica y el Desarrollo (OCED). El GBIF se fundó en marzo del 2001 y está abierto a la participación de cualquier país que lo desee, economía o organización reconocida internacionalmente que desee permitir el acceso a información sobre biodiversidad. La meta del GBIF es incorporar los datos de todos los niveles de biodiversidad que existen, desde las moléculas hasta los ecosistemas; sin embargo, en su primera fase se han enfocado a los niveles de información de especímenes y de especies. El GBIF tiene cuatro áreas de prioridad:

- La digitalización de los ejemplares.
- Acceso a esta información y conectividad entre bases de datos.
- Creación de catálogos electrónicos con los nombres científicos de las especies
- Desarrollo de nuevas metas y alcances.

El portal de Internet del GBIF, quien conecta todos los nodos participantes, fue puesto en línea en febrero de 2004 y para octubre de ese

mismo año ya contaba con 40 millones de registros (Canhos, Souza et al. 2004).

El GBIF ha trabajado junto con la iniciativa de Catalog of Life para tener una autoridad global para los aproximadamente 1.75 millones de nombres científicos que existen en el mundo, además de los sinónimos y nombres comunes que se usan para los organismos en distintas partes del mundo. El GBIF ha colaborando estrechamente con el CDB, proveyéndole toda la información de sus bases de datos. Dentro de su labor se encuentra también el financiamiento a organizaciones de todo el mundo para que digitalicen la información de sus colecciones, además de crear una red de expertos en informática de la biodiversidad y trabajar en la definición de asuntos como los derechos intelectuales de las colecciones (Canhos, Souza et al. 2004).

Por su parte en Europa se creó el Consortium of European Taxonomic Facilities (CETAF), la red europea más grande de información taxonómica, incluyendo información de museos de historia natural, jardines botánicos y otras colecciones biológicas. En su conjunto, el CETAF alberga una enorme cantidad de información zoológica, botánica, paleontológica y geológica, proveyendo recursos para el trabajo de un gran número de científicos de distintas disciplinas (CETAF 2005). El CETAF creó la red temática sobre fauna europea, que se complementa con el European Register of Marine Specimens (ERMS). En el área de botánica han desarrollado el proyecto European Flora y Euro-Mediterranean Plant Diversity.

En Europa existe también el European Natural History Specimen Information Network (ENHSIN), cuya meta es desarrollar una red interoperable y compartida de bases de datos de especímenes de colecciones biológicas de diversas instituciones europeas, donde han creado BioCASE, Biological Collection Access Service for Europe, un servicio de información basado en redes que permite consultas a las bases de datos de 35 instituciones en 30 países europeos además de Israel. Por otro lado, también en Europa se creó el ENBI (European Network for Biodiversity Information), que es la contribución europea al GBIF.

En América Latina, se ha desarrollado el Inter-American Biodiversity Information Network (IABIN), un foro basado en Internet para la cooperación técnica y científica que busca coordinar a los países americanos para

desarrollar, compartir y utilizar la información de la biodiversidad relevante para la toma de decisiones y la educación. Dentro de los proyectos de IABIN está una red para compartir información sobre especies invasoras de América, además de revisar ejemplares de colecciones europeas colectadas en América. IABIN está también desarrollando un acceso a la información contenida en las diferentes instituciones participantes. IABIN busca, a través del Global Environment Facility (GEF), crear una red de acceso a la información de la biodiversidad en América. Aquí se pretenden desarrollar también protocolos y estándares para acceder a la información acerca de la biodiversidad (Canhos, Souza et al. 2004).

En resumen, existe una gran cantidad de esfuerzos internacionales que buscan agrupar y difundir el conocimiento que se tiene acerca de la biodiversidad en el planeta, obtenida principalmente a partir de la información almacenada en las colecciones biológicas. Sin embargo, estos esfuerzos son inútiles si cada país no cuenta con la infraestructura adecuada para almacenar, clasificar y registrar los datos obtenidos de los ejemplares de las colecciones biológicas. Un primer paso para lograr el intercambio de información de las colecciones biológicas, es la creación de estándares internacionales.

#### **1.4 - Estándares internacionales**

El hombre ha sido consciente desde los albores de la civilización de la importancia de utilizar un mismo lenguaje, como queda plasmado en la historia bíblica de la Torre de Babel, en donde se narra cómo los hombres hace 4500 años fracasaron en la empresa de construir una torre que llegara hasta el cielo donde podrían vivir todos juntos sin tener que expandirse por toda la faz de la Tierra. Según se narra en un fragmento del Antiguo Testamento, Génesis 11, el dios de los hebreos, Yahveh quería que los hombres se expandieran sobre todos los continentes del mundo, y vio en esta obra una amenaza de sus deseos por lo que se propuso destruirla:

*“He aquí que todos forman un sólo pueblo y todos hablan una misma lengua, siendo éste el principio de sus empresas. Nada les impedirá que lleven a cabo todo lo que se propongan. Pues bien, descendamos y allí mismo confundamos su lenguaje de modo que no se entiendan los unos con los otros”*

Sembró la confusión entre los hombres que construían la torre haciendo que hablaran distintos idiomas, por lo que nunca pudieron entenderse entre sí haciendo que la torre se colapsara quedando en ruinas. Se le llamó la Torre de Babel, ya que Babel significa "confusión" en hebreo.

En el campo de la biología, el uso de distintos idiomas fue durante muchos siglos un problema para la clasificación de la vida debido a que los seres vivos recibían distintos nombres en distintos idiomas según las regiones. En el siglo XVII el médico sueco Carolus Linnaeus propuso un sistema de clasificación universal, la nomenclatura binomial, en donde el nombre científico de los seres vivos se compone del género y la especie, ambos escritos en latín o usando palabras latinizadas. El uso del latín es de gran importancia, ya que al ser una lengua muerta no puede tener los cambios naturales que sufre un idioma en uso. Así, los nuevos usos que se les dan a las colecciones biológicas hacen indispensable la creación de un sistema universal que permita la conjunción de distintas colecciones, ya sea de grupos diferentes como de instituciones y países distintos.

Los ejemplares de las colecciones biológicas se encuentran vinculados a información que nos permite identificar al organismo, además de explicar cómo, dónde, cuándo y por quién fue colectado (Beaman, Wieczorek et al. 2004). Este tipo de información suele encontrarse en todas las colecciones, lo que hace posible comparar ejemplares de distintas sedes. Esta práctica es muy común dentro de la comunidad científica, aunque poco eficiente, debido a que es necesario visitar físicamente diferentes colecciones ubicadas en regiones distantes. Por otro lado, existe el problema de la estandarización de la información en el que a pesar de que las distintas colecciones contengan datos semejantes, éstos no siempre se encuentran almacenados con los mismos formatos, lo que complica su compilación.

Los primeros esfuerzos realizados para estandarizar en el ámbito internacional los campos de las colecciones biológicas se llevaron a cabo en 1985, durante la primer reunión del Taxonomic Databases Working Group (TDWG), llevado a cabo en las instalaciones del Conservatoire et Jardin botaniques, de Génova, Suiza (TDWG 2003). Durante esta reunión se discutió la manera como deberían de ser tratados, dentro de las bases de datos de las colecciones, los campos referentes a los autores, los títulos de las revistas y

libros, los datos geográficos y los nombres de las familias, géneros y especies de los ejemplares. El objetivo era resolver los problemas de coordinación para el intercambio de datos (TDWG 1985).

El TDWG formó en el año 2000 el Task Group on Access to Biological Collection Data (KU), un subgrupo encargado de desarrollar, por un lado, un protocolo específico para efectuar la búsqueda y la recuperación de datos de diversas colecciones biológicas de manera simultánea, y por el otro, las especificaciones para estandarizar los datos contenidos en las mismas.

Lo primero dió lugar al Distributed Generic Information Retrieval (DiGIR), y lo segundo al Darwin Core, un estándar sencillo que utiliza un esquema plano para los datos de las colecciones biológicas. Este estándar se desarrolló con el objetivo de ser utilizado mientras se elaboraba el ABCD Schema, un estándar más complejo basado en un esquema jerárquico más robusto que el Darwin Core (TDWG 2005). La diferencia entre un esquema plano y uno jerárquico es que el primero mantiene todos sus elementos en el mismo nivel, mientras que en uno jerárquico se definen subgrupos.

El Darwin Core describe el formato que deben presentar 48 elementos (anexo 1) que representan los campos primarios de todas las colecciones biológicas. Esto no implica que las colecciones que utilicen este estándar se verán restringidas solamente a estos 48 elementos ya que éstas pueden contener los elementos que el curador de la colección considere, aún así podrán ser comparadas con otras colecciones de cualquier grupo, botánico o zoológico, a través de uno o varios de los 48 elementos que componen al Darwin Core. Hoy en día el Darwin Core se encuentra en la versión 2.0, y ha sido altamente aceptado como estándar internacional por la comunidad científica debido principalmente a dos factores: su sencillez y su compatibilidad con el DiGIR, de la cual hablaremos más adelante.

El 11 de diciembre de 2002 apareció finalmente la versión 1.01 del ABCD Schema, el estándar que buscaba desarrollar el TDWG. Es utilizado por primera vez en el 2003 por el proyecto Biological Collection Access Service for Europe (BioCASE), cuyo objetivo es crear una red europea de datos de colecciones biológicas. Este proyecto desarrolló su propio protocolo capaz de efectuar la búsqueda y la recuperación de datos de diversas colecciones biológicas de manera simultánea siempre y cuando utilicen el estándar

jerárquico ABCD Schema, ya que el DiGIR sólo puede trabajar con esquemas planos como el Darwin Core.

El ABCD Schema contiene 700 elementos estructurados en subgrupos, lo que le otorga un nivel de detalle muy alto, considerando de esta manera casi todos los campos contemplados en las distintas bases de datos de las colecciones biológicas. El ABCD Schema contiene también elementos de texto libre, mediante los cuales se pueden describir campos no contemplados por el ABCD esquema (TDWG 2005).

En 1999 nace el Global Biodiversity Information Facility (GBIF), a partir del Megascience Forum Working Group on Biological Informatics. Este proyecto busca crear una red internacional de bases de datos de colecciones biológicas. El GBIF reconoció tanto al Darwin Core como al ABCD Schema como estándares internacionales, permitiendo que coexistan dentro de su red. Para lograr ésto crearon un mapeo entre los dos estándares<sup>1</sup>, gracias al cual se pueden hacer equivalencias entre los campos propuestos por ambas iniciativas. No todos los elementos del ABCD Schema tienen una equivalencia en el Darwin Core, solamente los principales; y por el contrario, todos los elementos contemplados en el Darwin Core tienen una equivalencia con el ABCD Schema. Esto puede representar una ventaja o una desventaja para ambos, pero definitivamente ambos esquemas son válidos y de gran utilidad. Para definir cuál debe ser usado en una colección biológica deben analizarse varios factores antes de tomar una decisión.

En el caso de la propuesta de este trabajo, se recomienda utilizar el modelo del Darwin Core como estándar para aplicarse en muchas colecciones, ya que al ser mucho menos complejo que el ABCD Schema, elevaría el nivel de versatilidad del sistema, posibilitando que las metas propuestas de conectar todas las colecciones sea viable en un periodo de tiempo razonable. De esta manera, en el momento en el que una colección tenga los campos del Darwin Core listos para integrarse al sistema, podrá comenzar a ser consultada desde Internet y ser comparada con otras colecciones.

---

<sup>1</sup> <http://www.bgbm.org/TDWG/CODATA/Schema/Mappings/DwC2.0.htm>

Al utilizar el Darwin Core muchos campos específicos de las colecciones no quedarían contemplados en las consultas, cosa que sí sucede en el ABCD Schema.

Por otro lado, no hay que olvidar que ambos estándares sufren continuas modificaciones para adaptarse mejor a las necesidades de las colecciones, por lo que en el futuro podría llegar a ser conveniente el uso del ABCD Schema. Este esquema, sin embargo, requiere lograr reducir su nivel de complejidad sin perder su especificidad, cosa que probablemente alcance en algún momento. Mientras logra cumplir este objetivo, el ABCD Schema es un estándar muy inestable que sufrirá muchos cambios en el futuro (Hobern 2003).

En el caso de que el ABCD Schema lograra ser en el futuro una opción más viable que el Darwin Core, la migración del Darwin Core al primero se realiza de forma sencilla, por este motivo no es descabellado pensar en un futuro cambio de estándar.

## **1.5 - Protocolos de comunicación**

Una vez que los datos de una colección biológica se encuentran en un formato estandarizado como es el Darwin Core, se debe conectar a otras colecciones que compartan el mismo formato para que puedan ser consultadas de forma simultánea. Para lograr esto se debe utilizar un lenguaje común, llamado protocolo, entre ambas bases de datos. Los protocolos permiten la comunicación entre computadoras distintas, facilitando la transferencia de datos, como lo es por ejemplo, el Internet, que utiliza el protocolo HTTP o FTP.

En el caso de las colecciones biológicas el primer protocolo que se utilizó fue el Z39.50 para conectar distintas bases de datos de colecciones biológicas con el proyecto Species Analyst<sup>2</sup>. El Z39.50 es un estándar de comunicación desarrollado para solucionar los problemas de búsquedas en distintas bases de datos, presentando la enorme ventaja de que permite la conexión entre sistemas con distinto software y hardware. El Z39.50 está basado en un modelo cliente-servidor, en donde existe un origen y un destino. El origen se trata de la computadora del usuario, que envía la consulta que éste elabora a la computadora objetivo, que es la que maneja el sistema de bases

---

<sup>2</sup> <http://speciesanalyst.net/>

de datos. En el sistema origen se debe instalar un software local, el cual traduce las consultas al estándar Z39.50 y la envía al sistema objetivo o servidor.

Existen tres versiones del Z39.50, la versión 1 apareció el 1988, la 2 en 1992 y, finalmente, la tercera en 1995. Actualmente este protocolo es muy utilizado por bibliotecas virtuales, ya que presenta una gran cantidad de servicios especializados como lo es el mercado de la información con el estándar bibliográfico MARC21 (NISO 2003). Sin embargo, en el caso de las colecciones biológicas, el uso de este protocolo fue sustituido rápidamente por el Distributed Generic Information Retrieval (DiGIR) desarrollado al mismo tiempo que el estándar Darwin Core por el TDWG.

El DiGIR es un protocolo y una serie de herramientas que sirven para conectar un conjunto de bases de datos independientes entre si, y así poder consultarlas como si se tratara de una sola. Al igual que el Z39.50, el DiGIR está basado en un sistema cliente-servidor (Vieglais 2003). El DiGIR está conformado por una serie de proveedores de datos que envían la información de las bases de datos que lo utilizan a un portal mediante un protocolo establecido (Fig. 1). También incluye un registro en donde se encuentra almacenada la información de todos los proveedores de datos.

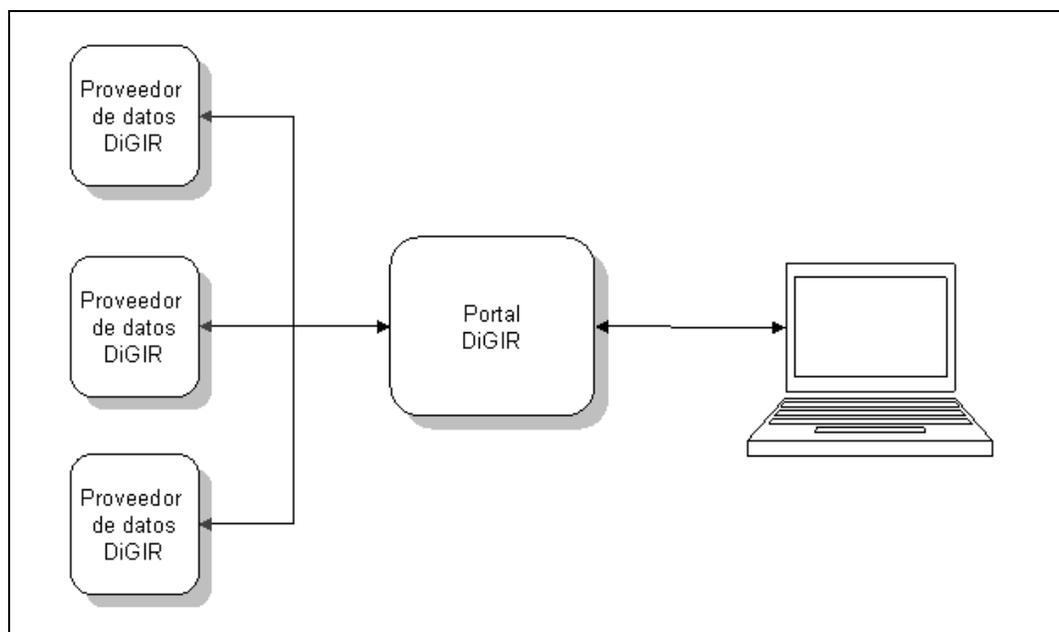


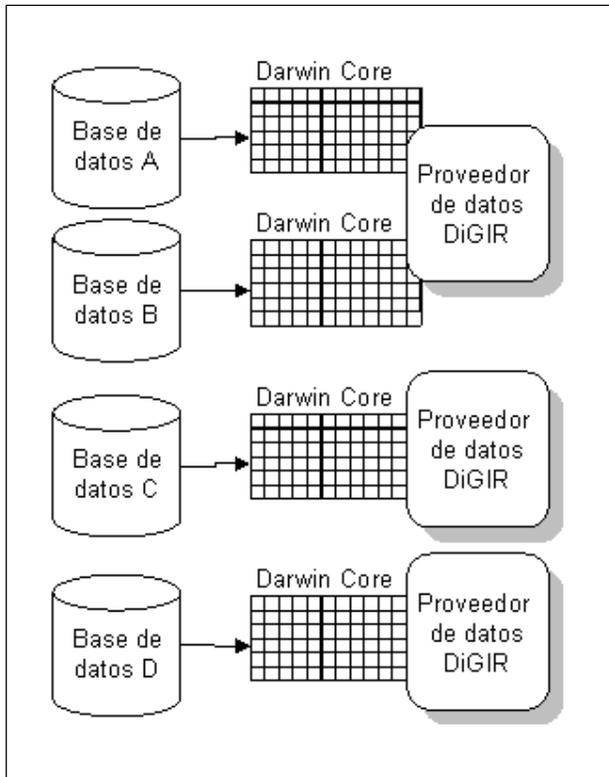
Figura 1. Sistema cliente-servidor desarrollado en el DiGIR.

El proveedor de datos debe poder comunicarse con el portal usando mensajes formateados de acuerdo al protocolo, tanto para recibir las consultas como para enviar las respuestas, o en el caso en que sucedan, mensajes de error. Toma la información de las bases de datos planas que utilizan el estándar del Darwin Core y sus extensiones, pudiendo albergar una o más bases de datos (Fig. 2). Debe ser capaz también de aceptar actualizaciones en sus datos y enviárselas al portal. Se recomienda que los proveedores de datos DiGIR estén dados de alta en el registro para que puedan ser ubicadas y utilizadas por varios portales.

El proveedor de datos se comunica utilizando el protocolo, el cual transmite los datos estructurados dentro de un XML Schema<sup>3</sup> que viaja por el Internet utilizando el protocolo HTTP. Dentro del protocolo debe encontrarse por lo menos la información necesaria para localizar al ejemplar: la dirección IP del proveedor de datos, el código de la institución y colección que alberga al ejemplar, el número de catálogo y la fecha de la última modificación de ésta información.

---

<sup>3</sup> El XML (eXtensible Markup Lenguaje) es un lenguaje de marcado que utiliza etiquetas definidas por el usuario para la transmisión y publicación de la información. El XML Schema es la estructura de las etiquetas definida también por el usuario. Tanto el Darwin Core como el ABCD Schema están desarrollados en un XML Schema que es transparente para el usuario.



*Figura 2. Los proveedores de datos DiGIR toman la información una o más bases de datos planas que utilizan el estándar internacional Darwin Core.*

El portal se comunica a través del protocolo con los proveedores de datos realizando exactamente la misma consulta a cada uno de ellos. Recibe también los mensajes de error enviados por los proveedores y les asigna un tiempo definido de respuesta. No sólo transmite los datos de los ejemplares sino también los datos de los proveedores, así como realiza consultas periódicas al registro para encontrar nuevos proveedores. Finalmente el portal debe de tener una interfase gráfica que permita que el usuario defina las consultas y a través de la cuál se publiquen los resultados de la misma. El portal DiGIR es el único módulo visible para los usuarios, el resto de los módulos no son visibles más que para el portal. Por último el registro es donde se almacena el nombre, la forma de acceso y los servicios de información que prestan los proveedores. El registro del DiGIR puede ser consultado por cualquier portal DiGIR instalado en el mundo (KU 2004).

## 1.6 - Bases de datos

La información que acompaña a los ejemplares de las colecciones biológicas, como por ejemplo el sitio donde fue colectado, la fecha, la identificación taxonómica etc., siempre ha sido considerada de enorme relevancia por los curadores. Siempre se ha prestado mucha atención en cómo guardar y preservar esta información para que pueda ser consultada y analizada posteriormente. Con el desarrollo a nivel mundial de la informática, se desarrolló también una herramienta que demostraría su utilidad, a tal grado, que hoy en día no sería posible concebir nuestra sociedad como la conocemos sin ella: las bases de datos. Fueron creadas en los años 60's y uno de sus principales objetivos fue el de administrar la inmensa cantidad de información que requería en el proyecto Apolo para enviar a un hombre por primera vez a la Luna. Posteriormente, las bases de datos serían utilizadas también en distintas áreas de la administración y del comercio, hasta que hoy en día son ampliamente usadas en el ámbito científico.

En el área de biología, se han desarrollado distintas bases de datos para administrar la información que acompaña a los ejemplares de las colecciones biológicas. Estas bases de datos por lo general pueden almacenar información taxonómica, curatorial, bibliográfica, geográfica, ecológica, etc. Dado el intenso uso que se hace hoy en día de las bases de datos, se han realizado esfuerzos por homogenizarlas, buscando crear sistemas de bases de datos distribuidas que superen los problemas de intercomunicación entre éstas.

Una base de datos distribuida es un conjunto de bases de datos relacionadas lógicamente que se localizan en distintas computadoras conectadas a través de una red de telecomunicaciones. Este conjunto de bases de datos debe estar administrado por un sistema de administración de bases de datos común, para asegurarse que lo que se realice en una base de datos no afecte al resto (Elmasri y Navathe 2003).

A continuación se presentan las características de algunas de las principales bases de datos desarrolladas para la administración de las colecciones biológicas, esfuerzos que han marcado la pauta para abrir nuevos campos en la investigación científica, como lo es la informática de la biodiversidad (Soberon y Peterson 2004).

### 1.6.1 – Specify

En 1996 se desarrolló el proyecto MUSE (KUNHM 1997) en el Museo de Historia Natural de la Universidad de Kansas, que buscaba proveer a los curadores de las colecciones biológicas de un *software* que les permitiera administrarlas. Posteriormente, y basados en la experiencia de MUSE desarrollaron un nuevo *software* llamado Specify que utiliza *Microsoft SQL Server 2000*<sup>4</sup>.

El diseño de la base de datos es sencillo y pueden realizarse operaciones sobre ésta a partir de comandos de SQL<sup>5</sup> como generar consultas para ordenar, filtrar y extraer datos de la base de datos. Specify es un *software* gratuito que contiene un proveedor de datos DiGIR y está autenticado por *Microsoft Windows*, lo cual asegura su correcto funcionamiento en el sistema operativo Windows. Incluye catálogos de especies, interfases personalizables y herramientas de información geográfica (Beach 2005). Su versatilidad se refleja en la popularidad de su uso en proyectos como MANIS<sup>6</sup>, HerpNet<sup>7</sup>, FishNet<sup>8</sup>, Lifemapper<sup>9</sup>, OBIS<sup>10</sup> y GBIF<sup>11</sup>.

Specify es un *software* muy bien desarrollado, sin embargo, al estar diseñado para un espectro tan amplio de usuarios (curadores de colecciones biológicas de todos los grupos taxonómicos) se vuelve un tanto complicado, principalmente en el momento de su instalación, en donde el mismo usuario debe definir qué campos son los que utilizará en las capturas que realice.

### 1.6.2 – Biota

Robert Colwell, de la Universidad de Connecticut, desarrolló otro *software* para la administración de las colecciones biológicas llamado Biota (Colwell 2005), que utiliza el administrador de bases de datos *4D*<sup>12</sup> y está certificado por *Mac OS*. Éste es un *software* comercial, distribuido por *Sinauer Associates Inc.*, que

---

<sup>4</sup> *Microsoft SQL Server 2000* es un administrador de bases de datos desarrollado por *Microsoft*.

<sup>5</sup> Structured Query Language, es un lenguaje estándar de comunicación con bases de datos.

<sup>6</sup> <http://elib.cs.berkeley.edu/manis/>

<sup>7</sup> <http://www.herpnet.org/documents/intro.html>

<sup>8</sup> <http://habanero.nhm.ku.edu/fishnet/>

<sup>9</sup> <http://lifemapper.org/>

<sup>10</sup> <http://www.iobis.org/>

<sup>11</sup> <http://www.gbif.net/portal/index.jsp>

<sup>12</sup> *4D* es un administrador de bases de datos desarrollado por Macintosh.

permite la exportación e importación de datos, además de su publicación en Internet. Facilita también el uso de imágenes y de referencias bibliográficas utilizando *EndNote*<sup>13</sup>. Existe en versiones para *Macintosh* y para *Windows* (Colwell 2004).

Biota es utilizado en proyectos como Palomillas de México<sup>14</sup>, y por colecciones como las del *The Mountain Lake Biological Station*<sup>15</sup>, y la colección entomológica de la Universidad de Colorado<sup>16</sup>. A pesar de sus ventajas en cuanto a la publicación electrónica, Biota presenta una gran desventaja: se trata de un *software* comercial, por lo que se deben comprar tanto el sistema como las actualizaciones que se desarrollen en el futuro.

### 1.6.3 – Atta

El Instituto Nacional de Biodiversidad (INBio) de Costa Rica, desarrolló su propio *software* llamado Atta, aunque hasta el momento sólo es para uso interno a la institución. Atta está construido por módulos para la generación, consulta, edición y análisis de información taxonómica, geográfica, ecológica y de usos potenciales de los ejemplares resguardados en sus colecciones biológicas (INBio). Presenta también un sofisticado sistema de publicación electrónica de especies llamado UBIs (unidades básicas de información)<sup>17</sup>, en donde se muestra información relacionada a la historia natural, demografía, conservación, distribución, información taxonómica y referencias bibliográficas de cada especie costarricense, además de presentar ilustraciones, tales como fotografías y dibujos de la especie. Cada UBI contiene también un mapa con la referencia geográfica de los sitios del país donde se recolectó la especie (INBio 2004). Una ventaja que presenta Atta es que la segunda generación de este tipo de administradores de datos desarrollado por la INBio (generado a partir de un sistema anterior llamado BIMS, creado en 1989).

Atta requiere de un equipo técnico que lo administre, ya que a pesar de que existe una base de datos común para todas las colecciones biológicas,

---

<sup>13</sup> *EndNote* es un software comercial desarrollado por *Thomson ResearchSoft*, que crea bases de datos bibliográficas, es muy utilizado por la comunidad académica para administrar las referencias bibliográficas de los artículos que producen.

<sup>14</sup> <http://cgic.ucol.mx/~mabl/indexe.htm>

<sup>15</sup> <http://www.mlbs.virginia.edu/BiotaWeb/collections.html>

<sup>16</sup> [http://cumuseum.colorado.edu/Research/Entomology/ento\\_databases.html](http://cumuseum.colorado.edu/Research/Entomology/ento_databases.html)

<sup>17</sup> <http://damis.inbio.ac.cr/ubis/default.htm>

cada una de ellas cuenta con una interfase de captura única, diseñada según las necesidades de la colección. Por este motivo es difícil que se convierta en un *software* de fácil distribución, como en el caso de los anteriores. Por otro lado, al estar desarrollado en *Oracle*, un manejador de bases de datos de carácter comercial, se vuelve un sistema muy caro de implementar, y hay que tomar en cuenta que muchas colecciones no cuentan con los recursos económicos necesarios para adquirir este tipo de desarrollos.

#### **1.6.4 – Biótica**

En 1995, la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), órgano gubernamental de México, desarrolló un sistema de bases de datos llamado Biótica, que homogeniza los datos de las colecciones biológicas pertenecientes al Sistema Nacional de Información sobre Biodiversidad (SNIB)<sup>18</sup>. Biótica está basado en *Microsoft Access*, aunque la última versión (4.1) permite que se migren los datos *Microsoft SQL Server 2000* (CONABIO 2003). Biótica está conformado por 70 tablas que se administran desde una interfase desarrollada en Visual Basic. Esta base de datos es de distribución gratuita<sup>19</sup>, sin embargo para activarlo es necesario contar con un nombre de usuario y una clave distribuidas por la CONABIO.

Para el desarrollo de Biótica se consideraron las necesidades de un gran número de profesionales como taxónomos, curadores, biogeógrafos, ecólogos, etnobiólogos etc. (CONABIO 2004) lo que la convierte en una herramienta muy poderosa capaz de ser distribuida y utilizada en un gran número de instituciones de distintas áreas. Cabe mencionar que Biótica es uno de los sistemas pioneros en la administración de colecciones biológicas a nivel mundial, gracias al cual México tiene hoy en día un desempeño relevante en el desarrollo de la informática de la biodiversidad.

Biótica es un sistema de bases de datos sumamente completo, pero a su vez complejo, ya que trata de abarcar todo el espectro de las colecciones biológicas de todos los grupos biológicos. Biótica es completo y complejo al mismo tiempo, esto se debe a la gran cantidad de tablas que contiene, lo que lo vuelve un tanto inestable. Tanto la captura de la información como su consulta

---

<sup>18</sup> <http://www.conabio.gob.mx/institucion/snib/doctos/acerca.html>

<sup>19</sup> [http://www.conabio.gob.mx/informacion/cgi-bin/biotica\\_http\\_biov4.1\\_esp.cgi](http://www.conabio.gob.mx/informacion/cgi-bin/biotica_http_biov4.1_esp.cgi)

resultan complejas, y estas últimas sumamente difíciles de realizar cuando se utilizan comandos de SQL directamente sobre la base de datos en Access. Estas complicaciones probablemente son la herencia de un sistema que se ha ido desarrollando sobre sus versiones anteriores desde 1996, capturando las necesidades de los distintos usuarios, pero a su vez almacenando problemas de diseño. Es importante comentar que, debido a la experiencia que ha adquirido la CONABIO con el desarrollo de Biótica, una segunda generación de este sistema sería probablemente el administrador de bases de datos de colecciones biológicas más eficiente de todas las opciones existentes.

## **1.7 - Biblioteca Digital**

Además de incluir ejemplares y datos sobre éstos, las colecciones biológicas contienen también algunas imágenes de los ejemplares y publicaciones que los describen. Estos documentos deben poder ser consultados también por Internet, en forma de objetos digitales (PDF, JPG, GIF, DOC, etc.).

Esta propuesta contempla la inclusión de dichos objetos digitales asociados a los ejemplares, pero también de imágenes, artículos, tesis, monografías, capítulos de libros, objetos multimedia etc., referentes a las especies, a los géneros y a las familias a las que los ejemplares pertenecen. De esta manera el usuario que consulte la base de datos de las colecciones, podría encontrar también información adicional sobre la especie o el género que esté consultando. Por lo general, los sistemas de administración de colecciones biológicas contemplan en su diseño de base de datos tablas donde se incluye la bibliografía y las imágenes asociadas. Esto implica que el usuario que realice una consulta sólo puede tener acceso a la información anexa que el curador decida incluir en la captura de los datos de cada ejemplar, por lo que se corre el riesgo de que esta información adicional no exista, o esté incompleta.

### **1.7.1 - Open Archives Initiative**

Como consecuencia de la reciente globalización y la llamada internacionalización del conocimiento, en los últimos años se ha modificado la forma de compartir la información científica. Hasta hace poco se encontraba confinada en las bibliotecas de todo el mundo, almacenadas en librerías o

estantes como “propiedad de la biblioteca”. Sin embargo, la necesidad que hoy en día se tiene de una distribución inmediata y eficiente de la información y el conocimiento ha derivado en la sustitución del formato impreso por el medio electrónico (i.e., publicaciones digitales) (Travieso 2003).

Con el creciente interés en la búsqueda de alternativas a los modelos tradicionales de comunicación científica, en muchas universidades del mundo se han desarrollado acervos de documentos electrónicos (repositorios), con el fin de mejorar la comunicación de los resultados de la investigación. A estos documentos se les conoce como *eprints*. Este concepto agrupa tanto a aquellos documentos que no han pasado por un proceso de certificación o revisión por pares (*preprints*) como aquellos que sí han pasado por tal proceso o *postprints* (artículos, libros, etc.). El más conocido de estos repositorios es arXiv.org<sup>20</sup> creado por Paul Ginsparg en Los Alamos (EU) para el área de Física.

En 1999 se llevó a cabo en Santa Fe, EU, una reunión que buscaba crear un sistema que permitiera que los acervos de *eprints* fueran compatibles entre sí. De esta manera nació Open Archives Initiative (OAI), que tenía como objetivo la creación de un protocolo capaz de recopilar los *eprints* de los distintos acervos con un sistema poco complejo. Este protocolo para la transmisión de contenidos en Internet se denominó OAI-PMH (*Open Archives Initiative – Protocol for Metadata Harvesting*) (Lagoze y Van de Sompel 2001).

En un principio OAI fue diseñado para intercambiar *eprints*, sin embargo, por su gran eficiencia, pronto fue adoptado por otros usuarios como museos y bibliotecas virtuales.

OAI, al igual que DiGIR, se compone principalmente de dos módulos: un proveedor de datos y un proveedor de servicios. El primero pone a disposición del segundo los metadatos<sup>21</sup> de los objetos a los que hace referencia (*eprints*, artículos, libros, imágenes etc.). Mientras que el segundo es un portal a través del cual el usuario puede realizar las consultas a los metadatos de los distintos proveedores (Figura 3). En pocas palabras, el proveedor de servicios define

---

<sup>20</sup> <http://arxiv.org/>

<sup>21</sup> Los metadatos son datos que describen otros datos, como por ejemplo el título de un libro, el tema, el autor etc. Es información sobre un objeto, más no la información del objeto en sí.

interfaces que facilitan la disponibilidad de contenidos procedentes de una variedad de distintos proveedores de datos.

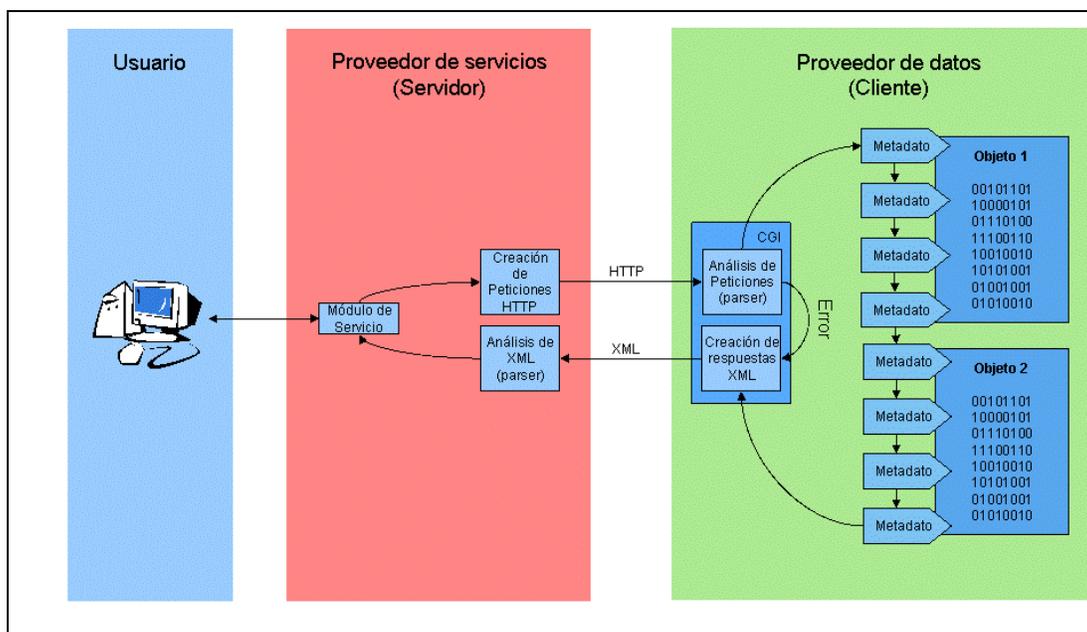


Figura 3. Consulta de metadatos, el usuario realiza una consulta que es enviada al CGI que es un sistema que analiza las peticiones para validarlas (en caso de que no sea posible devuelve un mensaje de error), el CGI realiza la búsqueda en los distintos metadatos de los objetos digitales para devolver las coincidencias usando un formato XML al usuario.

En resumen, podemos decir que OAI es un sistema dedicado a resolver problemas de interoperabilidad en bibliotecas digitales, definiendo protocolos simples aplicados al intercambio de metadatos.

Actualmente existen 255 proveedores de datos registrados en OAI (OAI 2005), un 40% más de los que había seis meses atrás, lo que indica la alta aceptación que ha tenido este protocolo a nivel mundial. Sin embargo, estas cifras están muy por debajo de la verdadera cantidad de proveedores de datos OAI que existen, ya que son pocos los que se registran en el sitio de OAI. Simplemente el proveedor de servicios AOIster<sup>22</sup> realiza consultas en 444 instituciones distintas, teniendo acceso a los metadatos de 5,195,319 objetos digitales.

### 1.7.2 - Dublín Core

<sup>22</sup> <http://oaister.umdl.umich.edu/o/oaister/>

OAI definió como estándar internacional de metadatos para identificar a los objetos digitales al Dublin Core<sup>23</sup> (ver capítulo 1), el cual contempla 15 metadatos que abarcan información sobre el contenido, la propiedad intelectual y el objeto digital (Tabla 1).

Tabla 1. Descripción de los metadatos del Dublin Core (RedIRIS 2003).

1. Título	Etiqueta: DC.Title	El nombre dado a un recurso, usualmente por el autor.
2. Autor o Creador	Etiqueta: DC.Creador	La persona u organización responsable de la creación del contenido intelectual del recurso. Por ejemplo, los autores en el caso de documentos escritos, artistas, fotógrafos e ilustradores en el caso de recursos visuales.
3. Claves	Etiqueta: DC.Subject	Los tópicos del recurso. Típicamente, Subject expresará las claves o frases que describen el título o el contenido del recurso. Se fomentará el uso de vocabularios controlados y de sistemas de clasificación formales. Se pueden ver ejemplos de vocabularios controlados en: Controlled vocabularies, thesauri and classification systems available in the WWW. DC Subject
4. Descripción	Etiqueta: DC.Description	Una descripción textual del recurso, tal como un resumen en el caso de un documento o una descripción del contenido en el caso de un documento visual.
5. Editor	Etiqueta: DC.Publisher	La entidad responsable de hacer que el recurso se encuentre disponible en la red en su formato actual, por ejemplo la empresa editora, un departamento universitario u otro tipo de organización.
6. Otros Colaboradores	Etiqueta: DC.Contributor	Una persona u organización que haya tenido una contribución intelectual significativa en la creación del recurso pero cuyas contribuciones son secundarias en comparación a las de las personas u organizaciones especificadas en el elemento Creator (por ejemplo, editor, ilustrador y traductor).
7. Fecha	Etiqueta: DC.Date	Una fecha en la que el recurso se puso a disposición del usuario en su forma actual. Esta fecha no ha de confundirse con la que pertenece al elemento Coverage, que sería asociada con el recurso sólo en la medida en que el contenido intelectual está de algún modo relacionado con esa fecha.

<sup>23</sup> <http://dublincore.org/>

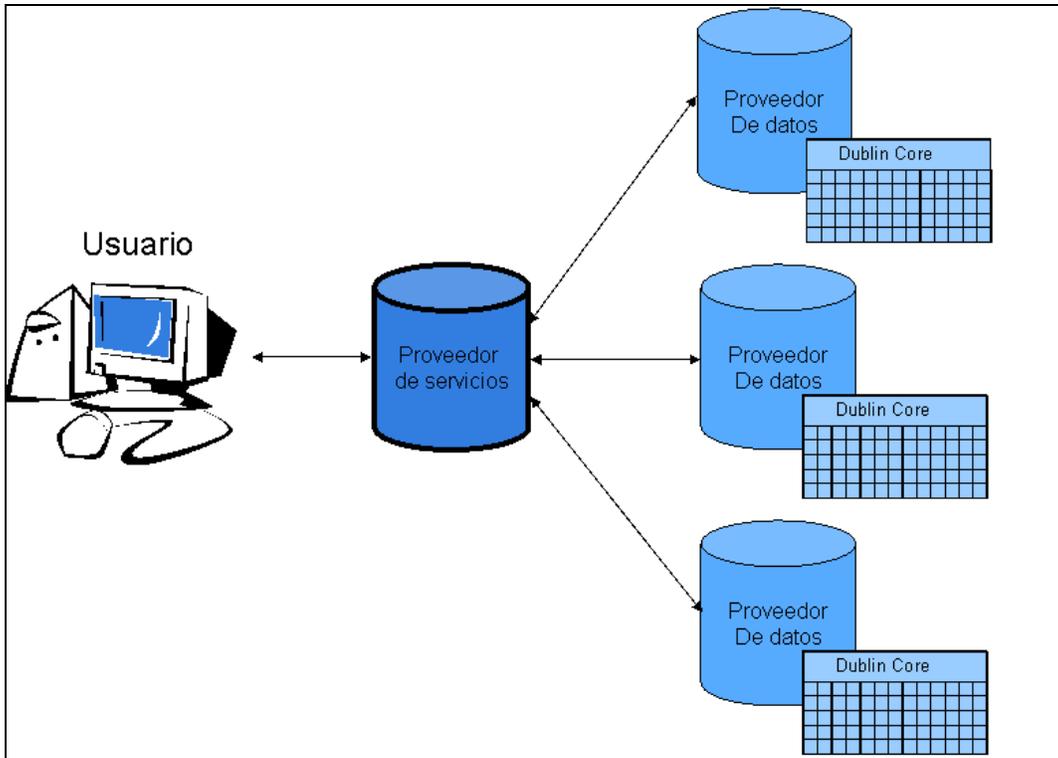
8. Tipo del Recurso	Etiqueta: DC.Type	La categoría del recurso, por ejemplo página personal, romance, poema, minuta, diccionario. Para asegurar la interoperabilidad, Type debería ser seleccionado de entre una lista de valores que actualmente se encuentra bajo desarrollo en un grupo de trabajo.
9. Formato	Etiqueta: DC.Format	El formato de datos de un recurso, usado para identificar el software y posiblemente, el hardware que se necesitaría para mostrar el recurso. Para asegurar la interoperabilidad, los valores de Format deberían ser seleccionados de entre una lista de valores que actualmente se encuentra bajo desarrollo en un grupo de trabajo.
10. Identificador del Recurso	Etiqueta: DC.Identifier	Secuencia de caracteres usados para identificar unívocamente un recurso. Ejemplos para recursos en línea pueden ser URLs y URNs (cuando estén implementados). Para otros recursos pueden ser usados otros formatos de identificadores, como por ejemplo ISBN ("International Standard Book Number" - Número Internacional Normalizado para Libros)
11. Fuente	Etiqueta: DC.Source	Secuencia de caracteres utilizado para identificar unívocamente un trabajo a partir del cual proviene el recurso actual. Por ejemplo, es posible usar Source con la fecha de 1603 como descripción de una película basada en una obra de Shakespeare, pero es preferible, en ese caso, usar Relation "IsBasedOn" con una referencia a un recurso distinto cuya descripción contenga el elemento Date con valor 1603.
12. Lengua	Etiqueta: DC.Language	Lengua/s del contenido intelectual del recurso. Prácticamente el contenido de este campo debería coincidir con los de la RFC 1766 (Tags para la identificación de lenguas, <a href="http://ds.internic.net/rfc/rfc1766.txt">http://ds.internic.net/rfc/rfc1766.txt</a> ); por ejemplo: en, es, de, fi, ja y zh.
13. Relación	Etiqueta: DC.Relation	Un identificador de un segundo recurso y su relación con el recurso actual. Este elemento permite enlazar los recursos relacionados y las descripciones de los recursos.
14. Cobertura	Etiqueta: DC.Coverage	La característica de cobertura espacial y/o temporal del contenido intelectual del recurso.
15. Derechos	Etiqueta: DC.Rights	Una referencia (URL, por ejemplo) para una nota sobre derechos de autor, para un servicio de gestión de derechos o para un servicio que dará información sobre términos y condiciones de acceso a un recurso. Una especificación

		formal del elemento Rights se encuentra actualmente en discusión y por lo tanto su uso se considera experimental.
--	--	---

Los metadatos del Dublin Core se presentaron para ser discutidos en una reunión de trabajo convocada por la *Online Computer Library Center* (OCLC), una organización sin fines de lucro que busca facilitar el acceso a la información mundial y reducir sus costos, y el *National Center for Supercomputing Applications* (NCSA) en el año de 1995 en Dublín, Ohio, Estados Unidos, de ahí se deriva su nombre. Fue diseñado tomando en cuenta la enorme proliferación de documentos digitales, y la imposibilidad de los bibliotecarios para catalogarlos.

El objetivo original del Dublin Core fue definir un conjunto de elementos que pudieran ser utilizados por los autores mismos para describir sus propios recursos en el Internet. Por este motivo se decidió que deberían de ser descriptores (originalmente 13) muy sencillos, que pudieran ser aplicados a cualquier objeto digital por el propio autor. Un proveedor de servicios de OAI realiza consultas a los metadatos de varios proveedores de datos OAI, específicamente a los 16 metadatos definidos en el Dublin Core para devolver al usuario los objetos digitales que busca (Fig. 4).

Cabe mencionar que OAI es un sistema que no requiere que los recursos sean gratuitos, o que forzosamente estén disponibles. Por lo tanto no es necesario que el objeto digital se encuentre dentro del proveedor de datos. Éste puede localizarse en cualquier otro servidor, o incluso puede no encontrarse disponible en Internet, pero si sus metadatos están disponibles en el proveedor de datos, su localización será mucho más eficiente.



*Figura 4. Proveedor de servicios y proveedores de datos de OAI. Existen varios proveedores de datos que contienen la información de los metadatos en formato Dublín Core, éstos están disponibles para el usuario a través de un proveedor de servicios único.*

## **2 - PROBLEMÁTICA**

Actualmente en el Instituto de Biología existen millones de ejemplares biológicos distribuidos en las distintas colecciones biológicas. Esto lo convierte en un referente muy importante para diversos estudios de biodiversidad que se realizan en el país y en el extranjero.

El acceso a esta información no siempre es sencillo o eficiente, ya que cada colección cuenta con bases de datos en formatos distintos. La única manera de realizar las consultas es en el lugar mismo donde se localiza la base de datos, a excepción de la Colección Nacional de Mamíferos. Ésta es la única colección del Instituto de Biología que se puede consultar de manera remota desde el sitio de MANIS.

### **3 - OBJETIVO**

El objetivo de este trabajo es proponer un sistema de información basado en Internet que permita la publicación electrónica de los datos de las colecciones biológicas almacenadas en el Instituto de Biología de la UNAM. Con esto se espera que nuestra universidad pueda participar de manera activa en las iniciativas nacionales e internacionales que buscan documentar la biodiversidad del planeta, además de apoyar a los investigadores y alumnos interesados en la biodiversidad. Por otro lado se buscará también asociar la información de las colecciones con objetos digitales vinculados a ésta, como pueden ser imágenes, artículos en formato digital etc.

#### **4 - METODOLOGÍA**

Con el fin de probar el estándar internacional Darwin Core y el sistema de cliente-servidor DiGIR, se migraron dos bases de datos de dos colecciones del Instituto de Biología a este estándar. Estas fueron la Colección Nacional de Anfibios y Reptiles y la Colección de Briofitas del Herbario Nacional. Ambas habían sido desarrolladas en Microsoft Access, una de ellas utilizando la interfase de Biótica.

Se seleccionaron ambas colecciones por tres motivos:

- 1) El número de ejemplares capturados en bases de datos de ambas colecciones es mayor a los 5000 ejemplares.
- 2) Ambas bases de datos se encuentran en distintos formatos, la CNAR está en Microsoft Access y la Colección de Briofitas utiliza el manejador de bases de datos Biótica, desarrollado por la CONABIO, que está construida en Microsoft Access.
- 3) Se escogió una colección zoológica y una botánica.

Con estos criterios se buscó realizar las pruebas con bases de datos representativas pero muy distintas entre ellas.

Posteriormente se instalaron dos proveedores de datos DiGIR para que albergaran a las colecciones una vez migradas. Por último se instaló también, un portal DiGIR desde el cual se pueden realizar consultas sobre ambas colecciones a través de Internet.

## **5 - RESULTADOS**

Se migraron ambas bases de datos al estándar Darwin Core. En el caso de la base de datos que se encontraba en el sistema Biótica la migración se llevó a cabo realizando consultas sobre la base de datos en Microsoft Acces. Los resultados de estas consultas fueron juntadas posteriormente en una tabla plana con los campos correspondientes al Darwin Core.

En el caso de la Colección que se encontraba en una tabla plana, simplemente se seleccionaron los campos que tuvieran un equivalente en el estándar Darwin Core y se copiaron en otra tabla plana con los campos del estándar Darwin Core. Sin embargo, en ambos campos hubo datos en las colecciones que no se pudieron migrar al estándar Darwin Core debido a que éste no los contempla, principalmente en el caso de la colección botánica. Por este motivo se recomienda utilizar además del estándar Darwin Core, extensiones que lo complementen.

### **5.1 - Extensiones al Darwin Core**

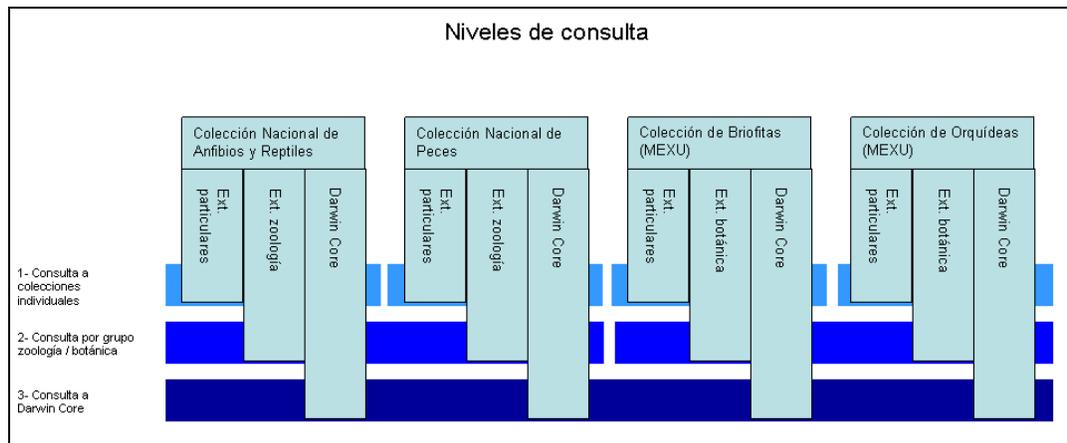
Para lograr que las colecciones biológicas del Instituto de Biología sean actores activos de las nuevas tendencias en sistemas de información biológica, es fundamental que adopten estándares internacionales, además que de esta manera se facilitaría una de las misiones de éste Instituto, que es la de difundir el conocimiento sobre la biodiversidad.

Al utilizar el Darwin Core como estándar para las bases de datos de las colecciones biológicas, se dejan fuera del sistema a muchos campos que existen tanto en las colecciones zoológicas como botánicas, como por ejemplo el tipo de agua (salobre o dulce) donde fue colectado un pez y la floración que presentaba una planta al momento de ser colectada. Esto se debe a que son campos que no comparten todas las colecciones, por lo que no pueden estar incluidos en el estándar Darwin Core.

Para resolver este problema se propone agregar campos adicionales al Darwin Core que contengan los valores que éste no contempla; a estos campos adicionales se les conoce como extensiones. A nivel mundial se han

desarrollado extensiones para distintos grupos como aves<sup>24</sup>, mamíferos<sup>25</sup>, microbios<sup>26</sup> y otras con campos curatoriales<sup>27</sup>.

En el caso del Instituto de Biología se propone utilizar extensiones con el fin de cubrir por completo todos los campos existentes en sus colecciones biológicas. La manera en la que se integrarán éstas al sistema es en tres niveles (Fig. 5).



*Figura 5. Existen tres niveles de profundidad en las consultas; en el nivel 1, la más específica, se podrán consultar todos los campos de una sola colección; en el nivel 2 se podrán consultar los campos de las extensiones de zoología/botánica y los campos contemplados por el Darwin Core de todas las colecciones que pertenezcan al mismo grupo (botánica o zoología), y en el nivel 3, la consulta más general, sólo se podrán consultar los campos del Darwin Core en todas las colecciones del IB.*

Por supuesto, en la medida en que sea posible, se propone utilizar las extensiones desarrolladas por proyectos internacionales para poder facilitar el intercambio de información. Finalmente, también pueden crearse extensiones más específicas que las de segundo nivel, correspondientes a ciertos campos creados para un proyecto en especial.

Este sistema de extensiones permitirá que las colecciones biológicas sean comparables en tres niveles, el primero y más general es a través de los campos del Darwin Core, en donde pueden analizarse todas las colecciones simultáneamente; en el segundo se podrán comparar todas las colecciones

<sup>24</sup> <http://iobis.org/FAQschema1.shtml>

<sup>25</sup> <http://bnhm.berkeley.museum/manis/DwC/darwin2jrw030315.xsd>

<sup>26</sup> <http://digir.net/schema/conceptual/darwin/extension/microbial/0.1/microbial.xsd>

<sup>27</sup> <http://digir.net/schema/conceptual/darwin/extension/curatorial/1.0/curatorialWithDiGIRv1.3.xsd>

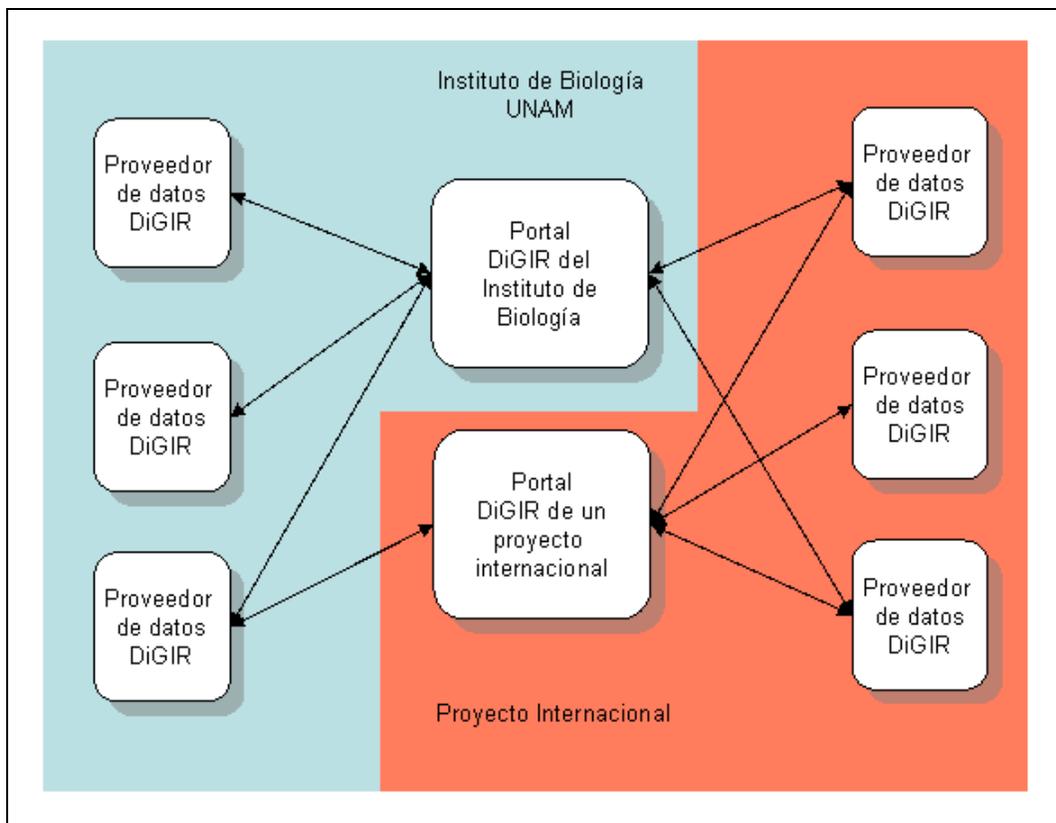
correspondientes a un área, ya sea botánica o zoológica; en el tercero y más específico, se podrán comparar sólo las colecciones que pertenezcan al mismo grupo, como por ejemplo anfibios y reptiles, briofitas etc. En este nivel de consulta se podrán comparar las colecciones del Instituto de Biología con colecciones de otras instituciones nacionales e internacionales siempre y cuando utilicen las mismas extensiones.

## **5.2 - El portal DiGIR en el Instituto de Biología**

En este trabajo se propone utilizar el DiGIR para poder acceder a la información almacenada en las colecciones biológicas del Instituto de Biología, ya que es un protocolo diseñado especialmente para este fin, además que funciona utilizando el estándar Darwin Core y acepta la anexión de extensiones. El DiGIR funciona en su totalidad a través de Internet, lo que permitirá que las colecciones del Instituto de Biología sean consultadas desde cualquier parte del mundo siempre y cuando se disponga de una computadora conectada a Internet. Otra ventaja de este protocolo es que es utilizado por muchos proyectos internacionales que cuentan con sus propios proveedores de datos<sup>28</sup>, lo que permitiría que desde el portal DiGIR del Instituto de Biología se puedan consultar también otros proveedores de datos. De la misma manera en que estos proyectos, a través de sus portales DiGIR podrán consultar los datos de los proveedores de datos del Instituto de Biología (Fig. 6). Esto le daría una cobertura mucho más amplia a las consultas realizadas desde el portal, aumentando también la visibilidad de los datos almacenados en las colecciones biológicas del Instituto de Biología.

---

<sup>28</sup> MANIS, FishBase, ORNIS, GBIF entre otros.



*Figura 6. El portal DiGIR del Instituto de Biología puede consultar tanto sus propios proveedores de datos como los pertenecientes a otros portales, de la misma manera que éstos pueden consultar también los proveedores de datos del Instituto de Biología.*

Por último, hay que mencionar que el DiGIR se está desarrollando como un sistema de código abierto<sup>29</sup>, lo que implica un ahorro ya que no debe comprarse ningún software, además de que en un futuro podría modificarse para que cumpla mejor las necesidades del Instituto de Biología.

Tanto el proveedor de datos como el portal DiGIR son software de distribución gratuita, por lo que para este trabajo se pudieron descargar de página de desarrollo del DiGIR<sup>30</sup>. En cada proveedor de datos se instaló una de las bases de datos con el estándar Darwin Core, y posteriormente se instaló el portal para configurarlo y poder consultar desde ahí las colecciones.

La interfase del portal DiGIR (Fig. 7 A) presenta tres secciones, en la primera se seleccionan la o las bases de datos que desean consultarse (Fig. 7 B), en la segunda se realiza la consulta seleccionando los campos del Darwin

<sup>29</sup> Los sistemas desarrollados bajo el concepto de código abierto son completamente gratuitos, además de que se puede modificar libremente el código fuente.

<sup>30</sup> [https://sourceforge.net/project/showfiles.php?group\\_id=38190](https://sourceforge.net/project/showfiles.php?group_id=38190)



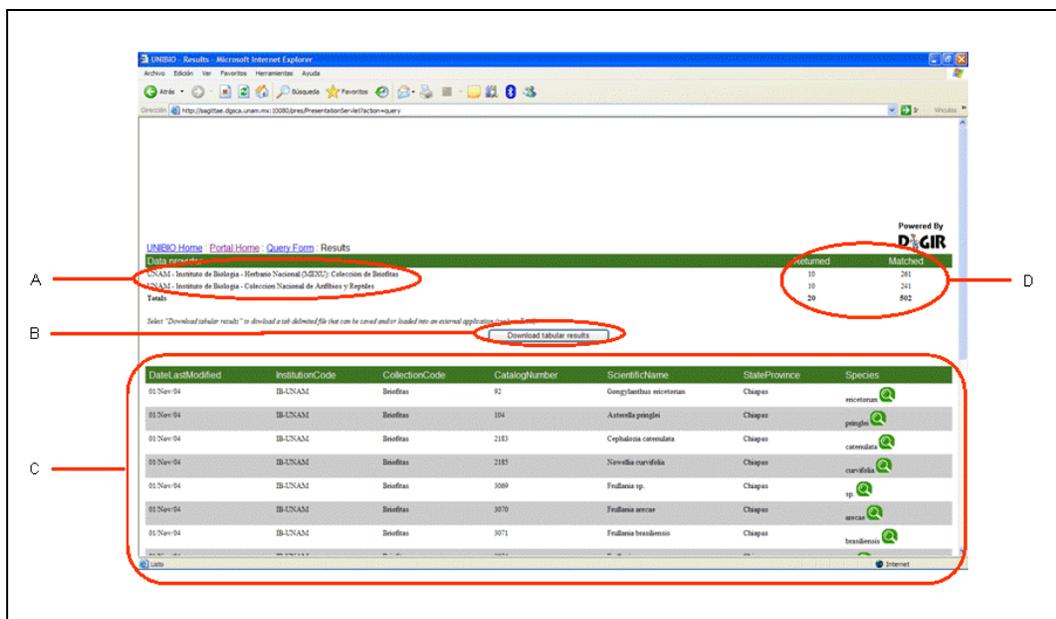


Figura 8. Vista de la interfase de resultados del portal DiGIR del Instituto de Biología.

De esta manera se demostró que el estándar Darwin Core y el portal DiGIR responden perfectamente a las necesidades del Instituto de Biología de la UNAM, ya que la consulta se realizó vía Internet y los resultados presentados corresponden con los que se encuentran en las bases de datos de ambas colecciones.

En una segunda etapa de pruebas se pretenden utilizar las extensiones tanto zoológicas y botánicas así como las particulares en el sistema.

### 5.3 - Una base de datos para el Instituto de Biología: recomendaciones

En el caso del Instituto de Biología se requiere aplicar un sistema que permita tanto la administración de las colecciones biológicas por parte de los curadores, como un registro y reporte actualizado de lo que se produce en el Instituto. Por ejemplo, una contabilidad total de los ejemplares y su estado en las colecciones, conocer los préstamos de ejemplares realizados y a quién fueron hechos, conocer el nivel de representatividad de sus colecciones a nivel de grupos taxonómicos y de zonas geográficas, con el fin de poder optimizar sus recursos al realizar nuevas colectas, etc.

El sistema que el instituto adopte debe facilitar tanto la consulta libre de sus ejemplares vía Internet (siempre y cuando el curador lo autorice), como la captura de los datos de ejemplares realizada por el personal especializado.

Para este fin se recomienda que ésta se realice también vía Internet, de manera tal que el personal encargado de esta misión no dependa de un *software* instalado en una computadora específica, sino que pueda realizar su labor desde cualquier computadora conectada a la Red Mundial. Esto facilita no solo la captura de información, sino también su revisión por los expertos, ya que no será necesario que éstos se encuentren en las instalaciones de cada colección para validar o modificar los datos. Al igual que en los casos de Atta y Specify, el instituto de Biología necesita una sistema en que las interfases de captura sean personalizadas para cada una de las colecciones, con el objetivo de que los campos que se presenten sean exclusivamente los que cada colección requiera.

Por otro lado, debido a que la captura y consulta de información se llevará a cabo por Internet, este sistema debe contar con un régimen de seguridad y validación de usuarios y datos muy sólido, para evitar falsas capturas o consultas no autorizadas. Para que los datos sean compatibles con otras bases de datos de proyectos nacionales e internacionales, el sistema debe poder migrar periódicamente sus datos a tablas con el estándar internacional Darwin Core, como fue revisado en el capítulo anterior.

Este sistema deberá ser capaz de otorgar a cada ejemplar de cada colección un identificador único, el cual pueda ser impreso en un código de barras para agilizar el movimiento tanto dentro como fuera del Instituto de Biología, eficientizando también las consultas *in situ* de los ejemplares. Además, debe permitir también la impresión de etiquetas en distintos formatos, para que de esta manera los curadores de las colecciones puedan renovar las etiquetas que identifican a los ejemplares preservados.

En el Instituto de Biología existen ejemplares colectados desde hace cien años, los cuales deben coexistir en el mismo sistema que los ejemplares de reciente adquisición. Esto representa un problema que debe ser resuelto. En muchos casos las etiquetas que identifican a los ejemplares más viejos contienen datos confusos y poco claros, principalmente en los campos referentes a nombres de colectores, determinadores y localidades. Estos nombres necesitan ser normalizados para que el sistema sea capaz de identificarlos y relacionarlos a los ejemplares más recientes, lo cual implica modificar la información original. Sin embargo, ésta también debe preservarse

de alguna manera, con el objetivo de que pueda ser verificada en el futuro. Para solucionar este problema, el sistema debe conservar la información original de las etiquetas, ya sea en una base de datos paralela o almacenando imágenes de las etiquetas originales. Este punto deberá ser discutido entre los curadores de las colecciones.

Otra de las características que debe presentar el sistema es la de contener catálogos con los nombres correctos o aceptados de los taxones de forma jerarquizada, para cada uno de los grupos de organismos de las colecciones. Esto permitiría tener un control en la calidad de los datos que ingresan al sistema. En este sentido, la CONABIO ha realizado una labor muy importante, concentrando catálogos de autoridades para Biótica<sup>31</sup>, gran parte de estos catálogos han sido elaborados por personal del Instituto de Biología, por lo que no debería ser difícil conseguirlos.

También es importante que existan catálogos de sinonimias, ya que organismos de la misma especie reciben nombres distintos en regiones distintas del país. Otro factor importante es que el sistema permita la actualización de la información taxonómica, dado que la taxonomía es un área muy versátil, que constantemente se modifica. Sin embargo, es importante mantener un registro de la información taxonómica usada para cada ejemplar en el pasado.

En muchos casos, es importante que la información del ejemplar esté acompañada por una imagen del mismo, principalmente en el caso de los ejemplares tipo, para que de ésta manera puedan identificarse ejemplares de forma remota. Las imágenes de los ejemplares tipo deben ser de muy alta resolución, con el fin de que puedan apreciarse los detalles, lo cual implica un problema si se desean poner en línea. Si las imágenes de los ejemplares de las colecciones están en línea, es muy importante que se puedan apreciar detalles para poder identificar a la especie a la que pertenecen. Para lograr esto se puede utilizar un servidor de imágenes, que permita visualizar detalles en imágenes de muy alta resolución, como por ejemplo FSI viewer<sup>32</sup>.

Idealmente, el registro de cada ejemplar también debería de ir acompañado por otros datos, como imágenes de ejemplares de la misma

---

<sup>31</sup> [http://www.conabio.gob.mx/informacion/catalogo\\_autoridades/doctos/electronicas.html](http://www.conabio.gob.mx/informacion/catalogo_autoridades/doctos/electronicas.html)

<sup>32</sup> <http://www.fsi-viewer.com/>

especie en su hábitat, referencias bibliográficas y artículos o textos completos asociados a la especie (o género) a la que pertenece el ejemplar en cuestión. En este caso es importante que la consulta de la información anexa a los ejemplares no quede reducida a las consultas que se realicen a la base de datos, sino que también puedan ser consultadas desde otros sistemas, como por ejemplo bibliotecas digitales. Con este fin se propone crear un sistema paralelo a la base de datos que relacione imágenes, referencias bibliográficas, artículos con los ejemplares de las colecciones.

Desafortunadamente no existe ningún sistema de bases de datos que cumpla completamente con estas características, para poder ser implementado en el Instituto de Biología. En este trabajo se propone aprovechar la experiencia adquirida en el desarrollo de las bases de datos Specfy, Atta y Biótica (Biota no puede ser usado debido a su carácter comercial) para crear un sistema de bases de datos diseñado especialmente para el IB. Se propone utilizar un modelo de bases de datos relacional utilizando en las tres bases de datos mencionadas, cuidando que se cumplan las necesidades de las colecciones biológicas del Instituto de Biología, como la impresión de códigos de barras para facilitar el manejo de los ejemplares de las colecciones, impresión de etiquetas en distintos formatos que contengan todos los datos de los ejemplares, administración de los préstamos de ejemplares para facilitar el intercambio de organismos entre el Instituto de Biología y otras instituciones académicas, migración a estándares internacionales como el Darwin Core para poder compartir la información almacenada en el Instituto con la de otras iniciativas a nivel internacional, control de la publicación de datos en Internet, ya que hay mucha información que los curadores no desean que se publique en Internet porque se trata de datos sobre especies en peligro de extinción, o datos que serán publicados próximamente. Por último, esta base de datos deberá poder ser migrada a Specify, Biota, Atta y Biótica, ya que hay muchos curadores e investigadores que las utilizan. Este diseño de base de datos permitiría albergar la información de todas las colecciones biológicas del Instituto de Biología.

Un sistema con las características anteriores tenderá a volverse complejo, sobretodo en su creación y administración, ya que requiere de una base de datos con muchos módulos, tanto para el almacenamiento de la

información como para su análisis. Para que este sistema sea viable, y devuelva más resultados que problemas, debe existir un área del instituto dedicada a su creación, actualización y administración (área de administración de información de colecciones), que se encargue de su desarrollo y de la resolución de problemas técnicos, de manera que los usuarios no tengan que lidiar con dichas complicaciones. Esta área del instituto tendría el objetivo de crear y administrar un sistema de bases de datos que sea capaz de extraer de manera eficiente toda la información almacenada en las bases de datos del Instituto de Biología y ponerla a disposición de los investigadores, docentes y público en general que ahí labora o que lo consulta frecuentemente. Para cumplir con este objetivo, esta área tendría que estar conformada por personal especializado tanto en biología como en informática. Esta figura de administración debe definir las tecnologías usadas y asegurarse de su interoperabilidad con centros nacionales internacionales. Para la selección de tecnologías es muy importante que se tomen en cuenta los sistemas ya desarrollados por otras instituciones.

Muchas colecciones biológicas del Instituto de Biología ya tienen adoptados sistemas como Biótica, Biota o Specify, y si sus usuarios encuentran estos sistemas funcionales para realizar sus labores no tiene sentido que migren su trabajo a un sistema nuevo. Por esta razón, el área de administración de información de colecciones deberá ofrecerles los mismos servicios que a los que decidan adoptar la nueva propuesta. Lo que sería importante es que todos los sistemas usados en el instituto sean compatibles entre sí, aprovechando los estándares internacionales desarrollados.

Esta propuesta busca recuperar lo que se ha desarrollado hasta el día de hoy en otras instituciones nacionales e internacionales, y adecuarlo a las necesidades particulares del Instituto de Biología, para optimizar la sistematización, el manejo y la difusión de la información que ahí se genera.

#### **5.4 - OAI en el Instituto de Biología**

En este trabajo se propone crear un proveedor de datos OAI dentro del Instituto de Biología, con el fin de que la información anexa a los ejemplares pueda ser recuperada no sólo cuando se consulten las colecciones biológicas sino desde

cualquier proveedor de servicios OAI en el mundo. Un proveedor de datos OAI aumentaría en gran medida la visibilidad de la información que posee el Instituto de Biología optimizando su misión de difusión y proporcionando la oportunidad de colaborar con otras instituciones.

La información que se aloje en el proveedor de datos deberá contar con los metadatos del Dublin Core, además de tres metadatos más que la asocien directamente con los ejemplares de las colecciones biológicas: familia, género y especie. De esta manera, cuando se consulte la base de datos de las colecciones biológicas a través del DiGIR, el usuario podrá consultar también la información anexa relacionada con la familia, el género o la especie del ejemplar consultado. Así mismo, los metadatos del Dublin Core permitirán que esta información esté disponible para los distintos proveedores de servicios OAI.

De la misma manera que la información de las colecciones es supervisada por los expertos en los distintos grupos, la información anexa que sea incluida en el proveedor de datos debe pasar por un proceso de revisión, con el fin de que éste presente datos relevantes. Por este motivo se propone la creación de un editor digital que revise los documentos y sus metadatos al momento de ser introducidos en el proveedor de datos. Utilizando el estándar Dublin Core, el proveedor de datos del Instituto de Biología cumpliría con los requisitos para ser un repositorio institucional, por lo que la información podría ser consultada por otras dependencias universitarias en todo el mundo. Este repositorio institucional podría estar asociado al sistema de Información basado en OAI <sup>4</sup>, que actualmente se está desarrollando en la Coordinación de Publicaciones Digitales de la UNAM (Galina 2004). Este sistema se construye a partir de una red de repositorios institucionales, proveedores de datos OAI.

Por otro lado, se propone también que el portal DiGIR funcione como proveedor de servicios OAI, lo que permitiría que en el momento de realizar consultas a los ejemplares de las colecciones biológicas, se pudieran realizar consultas en diversos proveedores de datos en todo el mundo. Este servicio implica un gran valor agregado al sistema, ya que se podría tener acceso a mucha información relevante relacionada con los ejemplares como por ejemplo artículos, imágenes, audio, videos, etc.

Las colecciones biológicas deben cumplir también la función de centros de información acerca de los grupos que trabajan, por lo que deben adquirir documentación especializada, o en su defecto conocer las referencias bibliográficas de ésta. Por este motivo se puede utilizar el proveedor de datos OAI también como un catálogo bibliográfico, en donde el usuario podrá investigar la bibliografía referente al grupo taxonómico que esté consultando.

## 6 - DISCUSIÓN

Ante la creciente necesidad de procesar eficientemente la gran cantidad de información que se ha almacenado y que se continúa produciendo en el Instituto de Biología de la UNAM y con el objetivo de ampliar el acervo mundial de datos que permitan entender mejor la biodiversidad de nuestro planeta, ha surgido la imperiosa necesidad de crear un sistema de digitalización, sistematización y publicación digital de la información de la biodiversidad que existe en el Instituto de Biología. En los capítulos anteriores se propusieron las bases para la creación de una red virtual de colecciones biológicas, construida a través de un sistema cliente-servidor llamado DiGIR, siguiendo estándares internacionales para la creación de bases de datos de colecciones biológicas como el Darwin Core. Esto permitiría acceder a la información de las colecciones biológicas del Instituto de Biología desde un portal de Internet. Para ampliar la efectividad y uso del sistema, esta red de bases de datos estaría conectada a su vez a una base de datos de publicaciones digitales, usando el sistema cliente servidor OAI-MHP, y aprovechando el estándar internacional para bases de datos de objetos digitales Dublín Core. Esto permitiría que, al mismo tiempo que se realicen consultas a las bases de datos de las colecciones biológicas, se pudieran realizar también consultas a objetos digitales asociados a los ejemplares de las colecciones tales como imágenes, artículos, audio, publicaciones y grupos de trabajo, entre otros.

Este sistema de bases de datos tiene la capacidad de poner al alcance de toda la comunidad científica del país y a nivel internacional, la información primaria producida en el Instituto de Biología de la UNAM. Gracias a este sistema el Instituto podría publicar la información de sus colecciones biológicas aumentando así su visibilidad, además de abrirse a la posibilidad de participar en diversos proyectos internacionales como lo son MANIS y el GBIF.

## 7 - ALCANCES A FUTURO

Los alcances de esta propuesta pueden escalarse a nivel de toda la UNAM, creando sistemas semejantes que sistematicen y publiquen de manera digital a través de Internet la información primaria producida en otros institutos de la universidad, tales como el Instituto de Geografía, el Instituto de Ciencias del Mar y Limnología, el Instituto de Geología etc.

Usando relativamente pocos recursos, podrían crearse distintas unidades de informática especializadas en distintas entidades de la UNAM con el mismo formato propuesto en este trabajo. Cada una de estas unidades de informática deberá digitalizar y sistematizar la información primaria producida en su dependencia cede. Por un lado, esto facilitaría el acceso a la información para los investigadores de la dependencia, y por el otro permitiría a la institución mantener un control sobre la información que ahí se produce. Se podrían realizar publicaciones electrónicas, análisis estadístico sobre la producción y también optimizar la asignación de recursos a partir del análisis de las áreas más desarrolladas.

Lo más importante es que estas unidades de informática tendrían que coordinarse entre si para asegurar la interconectividad de sus bases de datos, y de esta manera crear un sistema de información primaria interinstitucional, combinando datos acumulados en distintas dependencias, con lo que se podrían realizar consultas muy extensas. De esta manera podrían realizarse consultas a las bases de datos de varios institutos que resguardan mucha información muy relevante, de forma automática y a través de Internet,. Por ejemplo, podrían consultarse las bases de datos de las colecciones nacionales del Instituto de Biología, junto con las bases de datos del Instituto de Ciencias del Mar y Limnología, y junto con datos geográficos almacenados en el Instituto de Geografía. Incluso podrían empatarse estos datos científicos con datos almacenados en bases de datos de índole económica, social y bibliográfica.

Este sistema tendría que estar construido de forma modular, siendo cada instituto cede de un módulo muy similar al resto, con los mismos objetivos. Esto permitiría que el sistema funcione tanto con un solo módulo (e.g. bases de datos de un solo instituto) como con muchos (e.g. bases de

datos de varios institutos), ya que cada módulo sería independiente del resto, pero se complementarían con los demás. Un sistema modular haría que el proyecto arroje resultados prácticamente desde sus primeras etapas, desde la creación del primer módulo, y conforme se agreguen nuevos módulos los resultados irán aumentando.

Otra ventaja de construir un sistema modular es que cada uno de los módulos que forma parte del sistema completo también puede formar parte de otros proyectos. En el caso del módulo que se propone desarrollar en el Instituto de Biología, además de formar parte de esta red de unidades informáticas, también podría ser uno de los nodos del GBIF (Global Biodiversity Information Facility). De esta manera con un solo esfuerzo en la creación de cada módulo, se pueden cumplir varios objetivos.

Es muy importante que cada uno de estos módulos presente una estructura similar, capaz de asegurar la interconectividad con el resto. En un principio, los módulos deben consistir en migrar la información primaria que ya se encuentra en un formato digital, pero una vez que se haya completado este proceso, debe crearse un sistema de captura de datos, con el fin de mantener la información actualizada. En segundo lugar, cada módulo debe contar con un sistema de almacenamiento, que en todos los casos serán bases de datos. En este punto es importante remarcar que es muy importante realizar una profunda investigación sobre las tecnologías y los desarrollos informáticos existentes en el mundo. En el caso de las colecciones biológicas, es durante esta investigación donde se definió utilizar el Darwin Core y el DiGIR. En tercer lugar es fundamental asegurar que el sistema desarrollado o adquirido sea capaz de conectarse con el resto de los módulos de la red de unidades informáticas, pero también, en el caso en que existan, con otros proyectos tanto nacionales como internacionales. Por último, cada módulo debe tener una o varias interfases de consulta vía Internet, dependiendo de la cantidad y diversidad de información que manejen. En el caso del Instituto de Biología se propone tanto el desarrollo de un portal general, para consultar todas las colecciones biológicas, como el de una serie de portales más pequeños, sobre temas particulares como podría ser por ejemplo un portal de la Colección Nacional de Anfibios y Reptiles que obtenga sus datos del sistema general del

Instituto de Biología. De esta manera existiría una serie de portales divididos en tres niveles:

1. Portales sobre temas específicos (e.g. Portal de Anfibios y Reptiles). Estos portales se alimentan de los datos relacionados directamente con su tema, provenientes de las bases de datos de cada módulo.
2. Portales generales de cada módulo (e.g. Portal de las Colecciones Nacionales del Instituto de Biología). Esos portales se alimentan de todos los datos de un módulo.
3. Portal principal (e. g. Portal sobre biodiversidad y ambiente de México). Este portal se alimenta de todos los datos de todos los módulos de todas las entidades.

La administración de la red completa de unidades de informática sería relativamente sencilla, ya que basta con un sistema eficiente de comunicación entre los responsables de cada módulo, que puede estar basado en reuniones mensuales. No es necesario que exista un ente administrador de todo el sistema, ya que cada módulo es bastante similar a los otros.

Para la realización de este proyecto se requiere de una amplia labor multidisciplinaria e interinstitucional. Es necesario combinar conocimientos informáticos con las distintas especialidades de los distintos institutos participantes, como biología, geografía, oceanografía, geología etc.

## **8 - CONCLUSIONES**

En los últimos años se ha puesto en una situación crítica a la biodiversidad del planeta debido a la industrialización desmedida y a un sistema económico que contempla a la naturaleza como un recurso explotable bajo la premisa de mayor productividad a menor costo. Esto ha llevado a la necesidad de implementar iniciativas globales para revertir este proceso, asegurando la conservación de la biodiversidad del planeta junto con el desarrollo de las diversas sociedades que lo habitan.

Para lograr esto debe contarse con información muy completa de la biodiversidad en todo el planeta, además de sistemas que permitan ordenarla y analizarla, para así poder tomar las decisiones más apropiadas. En lo referente a la información sobre la biodiversidad, la humanidad cuenta con un enorme acervo que ha sido elaborado por el trabajo meticulado de colectores, taxónomos y naturalistas en los últimos trescientos años: las colecciones biológicas. Sin embargo, este acervo se encuentra disperso en todo el planeta, almacenado en diferentes instituciones de distintos países, lo que complica en gran medida su uso adecuado.

El desarrollo tecnológico de los últimos años, ha dado como resultado herramientas robustas de cómputo como las bases de datos y el Internet, las cuales se han planteado como la solución para lograr unificar el acervo mundial de información de biodiversidad representada en las colecciones biológicas de todo el mundo. Por otro lado, el cómputo también ha sido utilizado para analizar esa gran cantidad de información para comprender mejor los fenómenos que definen a la biodiversidad mundial. Un ejemplo de esto es el desarrollo de la informática de la biodiversidad, un área emergente de la ciencia que combina los adelantos tecnológicos de la computación con la gran cantidad de información almacenada en las colecciones biológicas, dando como resultado importantes estudios que son de gran importancia para la gestión de recursos naturales así como para las políticas de salud pública. Sin embargo, la informática de la biodiversidad sólo puede utilizar la información que está accesible en las colecciones biológicas: esto es, la información que se encuentra en formato digital, ordenada en bases de datos y accesible en la red.

El Instituto de Biología de la UNAM, además de ser la sede de las colecciones biológicas nacionales, tanto de zoología como de botánica, concentra también a los taxónomos más destacados en el país, lo que lo convierte en una pieza clave para la comprensión de la biodiversidad del país. Por esta razón, es fundamental que cuente con la infraestructura tecnológica que permita que su riquísimo acervo de información y sus especialistas puedan tomar parte dentro de la tendencia mundial que busca completar el rompecabezas de la biodiversidad.

Este trabajo representa una propuesta de cómo lograr este objetivo, utilizando estándares internacionales y sistemas probados con éxito en otras instituciones. Implementando estas tecnologías se espera que el gran acervo tanto de información como intelectual que se encuentra en el Instituto de Biología juegue un papel importante en la lucha mundial contra la pérdida de la biodiversidad.

## ANEXOS

### Anexo 1 - Campos definidos en el estándar internacional Darwin Core.

	Campo	Definición
1	DateLastModified	Última fecha en la que se hizo alguna modificación a la base de datos. AAAA-MM-DD hh:mm
2	InstitutionCode	Siglas que identifican a la Institución en donde se encuentra catalogada la colección
3	CollectionCode	Siglas que identifican a la Colección en la cual está el registro del organismo catalogado
4	CatalogNumber	Valor alfanumérico que identifica al registro dentro de la Colección
5	ScientificName	Nombre completo del nivel taxonómico inferior con el cual el organismo puede ser identificado
6	BasisOfRecord	Indica si el registro representa un objeto o una observación: observación, espécimen, organismo vivo...
7	Kingdom	Nombre del reino en el cual el organismo está clasificado
8	Phylum	Nombre del phylum o división en el cual el organismo está clasificado
9	Class	Nombre de la clase en la cual el organismo está clasificado
10	Order	Nombre del orden en el cual el organismo está clasificado
11	Family	Nombre de la familia en la cual el organismo está clasificado
12	Genus	Nombre del género en el cual el organismo está clasificado
13	Species	Nombre de la especie con el cual el organismo está clasificado
14	Subspecies	Nombre de la subespecie con el que el organismo está clasificado
15	ScientificNameAuthor	Nombre del autor del nombre científico
16	IdentifiedBy	Nombre de quien identificó al organismo
17	YearIdentified	Año de identificación
18	MonthIdentified	Mes de identificación
19	DayIdentified	Día de identificación
20	TypeStatus	Indica la clase del tipo nomenclatural que representa un espécimen
21	CollectorNumber	Número de identificación que establece las uniones entre partes/preparación de un espécimen y las notas de campo
22	FieldNumber	Número que identifica todo el material obtenido en un evento de colecta
23	Collector	Nombre(s) del colector o colectores del registro origina. Incluyendo ambos apellidos y las iniciales del nombre(s), se inicia con el apellido paterno
24	YearCollected	Cuatro dígitos que representen el año en el que se colectó el organismo
25	MonthCollected	Dos dígitos que representen el mes en el que se colectó el organismo
26	DayCollected	Dos dígitos que representen el día en el que se colectó el organismo
27	JulianDay	Día ordinal que le corresponde al día de colecta, dentro del año, se contabiliza desde el 31 de diciembre del año previo.
28	TimeOfDay	Hora del día en el que ocurrió la observación o la colecta en campo. Se expresa en formato de 24 horas y con decimales
29	ContinentOcean	Nombre completo del continente en donde se colectó el organismo
30	Country	Nombre completo del país o unidad política en donde se colectó el organismo
31	StateProvince	Nombre completo del estado, provincia o región en donde se colectó el organismo
32	County	Nombre completo del municipio en donde se colectó el organismo
33	Locality	Descripción de la localidad en donde se colectó el organismo. No debe haber información geográfica contenida en otros campos
34	Longitude	Longitud de la localidad en donde se capturó el organismo. Se expresa en grados decimales.
35	Latitude	Latitud de la localidad en donde se capturó el organismo. Se expresa en grados decimales.
36	CoordinatePrecision	Distancia, en metros, desde la latitud y longitud dadas hasta el límite más lejano de un círculo, imaginario, en el que esté inscrita la localidad completa. Cuando dicha distancia no se conozca se utilizará la palabra NULL

37	BoundingBox	Punto de acceso con una herramienta de búsqueda de información contenida en el proveedor de datos
38	MinimumElevation	Altitud mínima en metros, positiva si es sobre el nivel del mar y negativa si es por debajo del nivel del mar. Cuando en la base sólo exista una altitud se pondrá en altitud mínima
39	MaximumElevation	Altitud máxima en metros, positiva si es sobre el nivel del mar y negativa si es por debajo del nivel del mar. Cuando en la base sólo exista una altitud se pondrá en altitud mínima
40	MinimumDepth	Profundidad mínima en metros, dato proveniente de la cantidad de metros por debajo de la superficie del agua en donde se colectó el material. Se utilizan valores positivos para lugares por debajo de la superficie
41	MaximumDepth	Profundidad máxima en metros, dato proveniente de la cantidad de metros por debajo de la superficie del agua en donde se colectó el material. Se utilizan valores positivos para lugares por debajo de la superficie
42	Sex	Sexo del organismo
43	PreparationType	Piel, diapositiva, preservado
44	IndividualCount	Número de individuos de un lote
45	PreviousCatalogNumber	Número de catálogo anterior o proveniente de otra Institución
46	RelationshipType	Identifica el tipo de relación entre una colección y otra de referencia., p. e. Parásito de., epífita de...
47	RelatedCatalogItem	Referencia con otro espécimen (Código de la Institución, Código de la Colección, Número de Catálogo)
48	Notes	

## Anexo 2 - Resultados de una consulta hecha a la Colección Nacional de Anfibios y Reptiles utilizando el Portal DiGIR

DateLastModified	InstitutionCode	CollectionCode	CatalogNumber	ScientificName	BasisOfRecord	Kingdom	Phylum	Class	Order
9/27/2005 8:50	IBUNAM	CNAR	1	Aspidoscelis costata	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10	Aspidoscelis costata	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	100	Hyla eximia	Preservado	Animalia	Vertebrata	Amphibia	Anura
9/27/2005 8:50	IBUNAM	CNAR	1000	Sceloporus spinosus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10000	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10001	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10002	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10003	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10004	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10005	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10006	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10007	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10008	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10009	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	1001	Sceloporus spinosus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10010	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10011	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10012	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10013	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10014	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10015	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10016	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10017	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10018	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10019	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	1002	Sceloporus spinosus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10020	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10021	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10022	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria
9/27/2005 8:50	IBUNAM	CNAR	10023	Sceloporus grammicus	Preservado	Animalia	Vertebrata	Reptilia	Sauria

Family	Genus	Species	Subspecies	ScientificNameAuthor	IdentifiedBy	YearIdentified	MonthIdentified	DayIdentified
Teiidae	Aspidoscelis	costata	zweifeli	(Cope)				
Teiidae	Aspidoscelis	costata	zweifeli	(Cope)				
Hylidae	Hyla	eximia	wrightorum	Baird				
Phrynosomatidae	Sceloporus	spinosus	spinosus	Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann		9999		
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	spinosus	spinosus	Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann		9999		
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	spinosus	spinosus	Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				
Phrynosomatidae	Sceloporus	grammicus		Wiegmann				

Type	Status	CollectorNumber	FieldNumber	Collector	YearCollected	MonthCollected	DayCollected	JulianDay
				Zeferino Uribe Peña				
				A. RamirezE. PerezR. Martinez, A. Nieto				
				R. A. GuerreroJ.T. Medina	1932	6	8	
				Greg Mengden	1976	9	5	
				F. MendezM. Villagran	1989	7	19	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	8	22	
				F. MendezM. Villagran	1989	9	19	
				P. Ruthling	1975	9	13	
				F. MendezM. Villagran	1989	9	16	
				F. MendezM. Villagran	1989	9	16	
				F. MendezM. Villagran	1989	9	16	
				F. MendezM. Villagran	1989	9	16	
				F. MendezM. Villagran	1989	9	16	
				F. MendezM. Villagran	1989	9	16	
				F. MendezM. Villagran	1989	9	16	
				F. MendezM. Villagran	1989	10	18	
				F. MendezM. Villagran	1989	10	18	
				F. MendezM. Villagran	1989	10	18	
				V. V. Nogueron	1968	6	22	
				Fausto Mendez	1989	10	18	
				F. MendezJ. Mendez	1989	10	18	
				F. MendezH. Gadsden	1989	10	18	
				F. MendezH. Gadsden	1989	12	12	
				F. MendezH. Gadsden	1989	12	12	
				F. MendezH. Gadsden	1989	12	12	

TimeOfDay	Continent	Ocean	Country	StateProvince	County
			Mexico	Michoacan	Arteaga
			Mexico	Michoacan	Huacana
			Mexico	Estado de Mexico	San Felipe Del Progreso
			Mexico	Hidalgo	Zacualtipan De Angeles
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Puebla	Puebla
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Hidalgo	Atotonilco El Grande
			Mexico	Distrito Federal	Magdalena Contreras
			Mexico	Distrito Federal	Magdalena Contreras
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco
			Mexico	Estado de Mexico	Tlanguistenco

Locality	Longitude	Latitude	CoordinatePrecision
carretera Apatzingan-Playa Azul Rancho ,El Reparito, Km. 218			
Cupuan del Rio			
Potrero de Sn. Diego, 12 Km por carr. al W de Villa Victoria, Mpio. Villa Victoria, Edo. de Mex.			
Metzquititlan, 17.6 km N			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Los Alamos, S			
Monte Alegre, area Los Gavilanes			
Monte Alegre, area Los Gavilanes			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			
Capulin (entronque) 8 kms al S. del Ajusco			

MinimumElevation	MaximumElevation	MinimumDepth	MaximumDepth	Sex	PreparationType	IndividualCount	PreviousCatalogNumber
						1	
						1	
						1	
						1	
				hembra		1	
				macho		1	
				macho		1	
				hembra		1	
				hembra		1	
				macho		1	
				macho		1	
				hembra		1	
						1	
				hembra		1	
				hembra		1	
				hembra		1	
				macho		1	
				macho		1	
				macho		1	
				macho		1	
				macho		1	
				macho		1	
				hembra		1	
						3	
				hembra		1	
				hembra		1	
				hembra		1	
				macho		1	
				macho		1	
				macho		1	

## REFERENCIAS

- Beach, J. (2005). "Specify, [En línea], <http://www.specifysoftware.org/Specify/> (página consultada el 20 de febrero de 2005)."
- Beaman, R., J. Wiczorek, et al. (2004). "Determining Space from Place for Natural History Collections." D-Lib Magazine **10**(5).
- Berg, W., A. Johnels, et al. (1966). "Mercury content in feathers of swedish birds from the past 100 years." Oikos **17**(1): 71 - 83.
- Brailovsky, H. y B. Gómez Varela (1993). Colecciones Biológicas Nacionales del Instituto de Biología, Colecciones Zoológicas. México D. F., Instituto de Biología.
- Bush, V. (1945). "As We May Think." The Atlantic Monthly.
- Canhos, V. P., S. Souza, et al. (2004). "Global biodiversity informatics: setting the scene for a "new world" of ecological modeling." Biodiversity Informatics **1**: 1 - 13.
- CETAF (2005). "CETAF - Consortium of European Taxonomic Facilities, [En línea], <http://www.cetaf.org/index.php> (página consultada el 26 de febrero de 2005)."
- Colwell, R. K. (2004). "Biota: The Biodiversity Database Manager, [En línea], [http://viceroy.eeb.uconn.edu/Biota2Pages/about\\_biota.html#AboutFeatures](http://viceroy.eeb.uconn.edu/Biota2Pages/about_biota.html#AboutFeatures)(página consultada el 20 de febrero de 2005)."
- Colwell, R. K. (2005). "Robert K. Colwell Homepage, [En línea], <http://viceroy.eeb.uconn.edu/Colwell> (página consultada el 20 de febrero de 2005)." The University of Connecticut.
- CONABIO (2003). "Historia de Biótica, [En línea], [http://www.conabio.gob.mx/informacion/biotica\\_espanol/doctos/historia.html](http://www.conabio.gob.mx/informacion/biotica_espanol/doctos/historia.html) (página consultada el 26 de febrero de 2005)." Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.
- CONABIO (2004). "Historia de Biótica, [En línea], [http://www.conabio.gob.mx/informacion/biotica\\_espanol/doctos/acerca\\_biotica.html](http://www.conabio.gob.mx/informacion/biotica_espanol/doctos/acerca_biotica.html) (página consultada el 26 de febrero de 2005)." Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.
- Costa, J., A. T. Peterson, et al. (2002). "Ecologic niche modeling and differentiation of populations of *Triatoma brasiliensis* neiva, 1911, the most important Chagas' disease vector in northeastern Brazil (Hemiptera, Reduviidae, Triatominae)." American Journal of Tropical Medicine and Hygiene **67**(5): 516-520.

- Dávila Aranda, P. D. y M. T. Germán Ramírez (1991). Colecciones Biológicas Nacionales del Instituto de Biología, Herbario Nacional de México. México D. F., Instituto de Biología, UNAM.
- Ellegren, H., G. Lindgren, et al. (1997). "Fitness loss and germline mutations in barn swallows breeding in Chernobyl." Nature **389**: 593 - 596.
- Elmasri, R. y S. B. Navathe (2003). Fundamentals of Database Systems. Redwood City, CA, USA, Pearson.
- Fonseca, D. M., S. Campbell, et al. (2001). "Aedes (Finlaya) japonicus (Diptera: Culicidae), a Newly Recognized Mosquito in the United States: Analyses of Genetic Variation in the United States and Putative Source Populations." Journal of Medical Entomology **38**(2): 135 - 145.
- Galina, I. (2004). Sistema de Información Digital UNAM (Comunicación personal). México, Coordinación de Publicaciones Digitales, DGSCA.
- Hobern, D. (2003). Strategy for exchange of specimen and observation data. GBIF Data Access and Database Interoperability, GBIF.
- INBio "Atta, [En línea], <http://atta.inbio.ac.cr/atta03.html> (página consultada el 20 de febrero de 2005)."
- INBio (2004). "Especies de Costa Rica, [En línea], [http://darnis.inbio.ac.cr/ubis/que\\_son.html#que](http://darnis.inbio.ac.cr/ubis/que_son.html#que) (página consultada el 20 de febrero de 2005)."
- Krishtalka, L. y P. S. Humphrey (2000). "Can natural history museums capture the future?" Bioscience **50**(7): 611-617.
- KU (2004). "DiGIR, [En línea], <http://bgbm3.bgbm.fu-berlin.de/TDWG/CODATA/ABCD-Evolution.htm> (página consultada el 23 de enero de 2005)." Biodiversity Research Center Informatics.
- KUNHM (1997). "The KUNHM MUSE Project, [En línea], <http://www.biodiversity.uno.edu/muse/> (página consultada el 20 de febrero de 2005)." University of Kansas Natural History Museum.
- Lagoze, C. y H. Van de Sompel (2001). The open archives initiative: building a low-barrier interoperability framework, [Artículo en línea], <http://www.openarchives.org/documents/jcdl2001-oai.pdf> (artículo consultado el 12 de marzo de 2005).
- Lens, L., S. Van Dongen, et al. (2002). "Avian Persistence in Fragmented Rainforest." Science **298**(5596): 1236 - 1238.
- Martinez-Meyer, E., A. T. Peterson, et al. (2004). "Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene

extinctions and climate change projections for biodiversity." Global Ecology and Biogeography **13**(4): 305-314.

Monteiro, L. R. y R. W. Furness (1997). "Accelerated increase in mercury contamination in North Atlantic mesopelagic food chains as indicated by time series of seabird feathers." Environmental Toxicology and Chemistry **16**(12): 2489 - 2493.

NISO (2003). "Information Retrieval (Z39.50): application service definition and protocol specification." National Information Standards Organization Press.

OAI (2005). "Registered Data Providers, [Página en línea], <http://www.openarchives.org/Register/BrowseSites> (página consultada el 12 de marzo de 2005)." Open Archives Initiative.

Peterson, A. T., E. Martinez-Meyer, et al. (2004). "Modeled climate change effects on distributions of Canadian butterfly species." Canadian Journal of Zoology **82**(6): 851-858.

Peterson, A. T., M. A. Ortega-Huerta, et al. (2002). "Future projections for Mexican faunas under global climate change scenarios." Nature (London) **416**(6881): 626-629.

Peterson, A. T., V. Sanchez-Cordero, et al. (2002). "Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico." Emerging Infectious Diseases **8**(7): 662-667.

Peterson, A. T. y J. J. Shaw (2003). "Lutzomyia vectors for cutaneous leishmaniasis in southern Brazil: Ecological niche models, predicted geographic distributions, and climate change effects." International Journal of Parasitology(33): 919 / 931.

Peterson, A. T. y M. F. Siqueira (2003). "Consequences of global climate change for geographic distributions of cerrado tree species." Biota Neotropica **3**(2): 2 - 14.

Peterson, A. T. y J. Soberón (2004). "Biodiversity informatics: managing and applying primary biodiversity data." Philosophical Transactions of the Royal Society of London B **359**: 689 - 698.

Raxworthy, C. J., E. Martinez-Meyer, et al. (2003). "Predicting distributions of known and unknown reptile species in Madagascar." Nature **426**(6968): 837-841.

RedIRIS (2003). "Metainformación - Dublin Core, [Página en línea], <http://www.rediris.es/metadata/#Title> (página consultada el 12 de marzo de 2005)." Plan Nacional de Investigación y Desarrollo.

Sanchez-Cordero, V. y E. Martinez-Meyer (2000). "Museum specimen data predict crop damage by tropical rodents." Proceedings of the National Academy of Sciences of the United States of America **97**(13): 7074-7077.

Soberon, J., J. Golubov, et al. (2001). "The importance of Opuntia in Mexico and routes of invasion and impact of Cactoblastis cactorum (Lepidoptera : Pyralidae)." Florida Entomologist **84**(4): 486-492.

Soberon, J. y A. T. Peterson (2004). "Biodiversity informatics: managing and applying primary biodiversity data." Philosophical Transactions of the Royal Society of London B Biological Sciences **359**(1444): 689-698.

Suarez, A. V. y N. D. Tsutsui (2004). "The value of museum collections for research and society." Bioscience **54**(1): 66 - 74.

Taubenberger, J. K., A. H. Reid, et al. (1997). "Initial genetic characterization of the 1918 "Spanish" Influenza virus." Science **275**: 1793 - 1796.

TDWG (1985). Minutes of the first meeting. Geneva, Conservatoire et Jardin botaniques.

TDWG (2003). "Taxonomic Databases Working Group, [En línea], <http://www.tdwg.org/dates.html> (página consultada el 23 de enero de 2005)."

TDWG (2003). "Taxonomic Databases Working Group, [En línea], <http://www.tdwg.org/dates.html> (página consultada el 23 de enero de 2005)."

TDWG (2005). "ABCD Schema Design Principles, [En línea], <http://bgbm3.bgbm.fu-berlin.de/TDWG/CODATA/ABCD-Evolution.htm> (página consultada el 23 de enero de 2005)."

TDWG (2005). "Origin and Evolution of ABCD, [En línea], <http://bgbm3.bgbm.fu-berlin.de/TDWG/CODATA/ABCD-Evolution.htm> (página consultada el 23 de enero de 2005)

".

Thomas, C. D., A. Cameron, et al. (2004). "Extinction risk from climate change." Nature **427**: 145 - 148.

Travieso, M. (2003). "Las publicaciones electrónicas: una revolución en el siglo XXI." Acimed **11**(2).

Vieglais, D. (2003). "DiGIR Provider Manual, [En línea], [http://digir.net/prov/prov\\_manual.html](http://digir.net/prov/prov_manual.html) (página consultada el 30 de enero de 2005)."

Winker, K. (2004). "Natural history museums in a postbiodiversity era." BioScience **54**(5): 455 - 459.

Yates, T. L., J. N. Mills, et al. (2002). "The Ecology and Evolutionary History of an Emergent Disease: Hantavirus Pulmonary Syndrome." BioScience **52**(11): 989 - 998.