

Arquitectura de IABIN e interoperabilidad
Versión 1.1
Autor: Boris Ramírez, Secretaria de IABIN

1.	Introducción	2
2.	Conceptos Generales.....	2
a.	Conceptos Básicos para intercambiar datos.....	3
b.	Arquitectura Básica de IABIN.....	6
3.	Implementación de la Arquitectura Básica en cada una de las Redes	8
a.	Red Temática de Especies y Especímenes.....	8
b.	Red Temática de Especies Invasoras	9
c.	Red Temática de Polinizadores.....	10
d.	Red Temática de Ecosistemas.....	11
e.	Red Temática de Áreas Protegidas	13
f.	Red Geo-Espacial.....	14
g.	Catálogo	15
4.	Integración de las Redes	16
a.	Concepto Básico de Integración.	16
b.	Casos de uso en la Primera Fase (Proyecto IABN-GEF)	20
c.	Integración con Otra Iniciativas.....	20
5.	Estándares y Protocolos	21
6.	Consideraciones Generales de la Arquitectura de IABIN.	21
a.	Calidad de datos.....	21
7.	Anexos	23
	Anexo 1: Principios de IABIN.....	23
	Anexo 2. Diferencia entre IABIN Biobot y Google	24

1. Introducción

La Red Interamericana de Información sobre Biodiversidad (IABIN), fue concebida como un forum que propicia la colaboración técnica y la coordinación entre los países de las Américas para recolectar, compartir y utilizar información sobre biodiversidad que sea relevante para los procesos de toma de decisiones sobre conservación y manejo de la biodiversidad, así como para la educación en la región.

Bajo este objetivo se establecieron las siguientes metas:

- Construir una infraestructura para el intercambio de información digital sobre biodiversidad
- Fortalecer la capacidad técnica para intercambiar información entre los países de las Américas sin límites de fronteras políticas, lingüísticas e institucionales
- Proveer acceso a información sobre la biodiversidad que sea útil a los tomadores de decisiones para mejorar los esfuerzos de conservación de ésta.
- Mejorar la capacidad para almacenar, utilizar y distribuir información sobre biodiversidad que sea científicamente confiable y actualizada
- Producir o adaptar herramientas para la toma de decisiones en materia ambiental que propicien el desarrollo sostenible de la región.

Desde el 2005 la Secretaria de IABIN, en colaboración con cada una de las Instituciones Coordinadoras de las Redes Temáticas, ha liderado el proceso para alcanzar las metas y los objetivos establecidos inicialmente.

Este documento presenta el estado actual de implementación de cada una de las redes de IABIN, cómo estas redes se van a integrar inicialmente así como la integración de la información de los proveedores de datos conectados a IABIN y a otras iniciativas a nivel nacional, regional e internacional.

En todo este proceso se ha contado con la colaboración de varias organizaciones las cuales ya han venido trabajando en este tema y con las cuales IABIN ha establecido acuerdos, con la finalidad de adaptar sus tecnologías a las necesidades de IABIN e impulsar sus estándares y protocolos.

2. Conceptos Generales

Este capítulo tiene por finalidad definir algunos conceptos generales y la terminología utilizada en la implementación de la Red y la usada a través del presente documento.

Es importante aclarar que IABIN es una red descentralizada de datos en la cual los proveedores de datos son el componente esencial y la razón de ser de la red. Sin proveedores de datos, IABIN no podría existir.

a. Conceptos Básicos para intercambiar datos

Para alcanzar los objetivos generales de IABIN se hace necesario poner a disposición los datos sobre biodiversidad que cada organización posee. Estas organizaciones se convertirán en proveedores de datos de IABIN al momento que compartan sus datos con los demás miembros de la red siguiendo los principios establecidos por IABIN, para tal fin (ver anexo 1: Principios de IABIN).

Gran parte del éxito de la Red depende de la capacidad que posea la misma y sus colaboradores (Puntos Focales e Instituciones Coordinadoras) de conectar proveedores con datos de alta calidad y que estos datos puedan ser consultados de una manera rápida, eficiente y en tiempo real, y que al mismo tiempo el proveedor de datos siga teniendo el control sobre ellos y los siga manteniendo.

El proveedor de datos debe tener la discrecionalidad de decidir qué datos desea compartir, de qué forma los desea compartir y con qué redes desea compartir sus datos

Pero para poder darles esta discrecionalidad a los proveedores de datos, IABIN tuvo que resolver las siguientes tres preguntas.

- ¿Cómo integrar bases de datos que contienen datos de biodiversidad y que se encuentran en diferentes organizaciones y formatos?
- ¿Cómo hacer disponible la información de un proveedor de datos en tiempo real?
- ¿Cómo saber dónde se encuentran los datos?
- ¿Cómo hacer disponible la información que no está digitalizada?

-¿Cómo integrar bases de datos que contienen datos de biodiversidad y que se encuentran en diferentes organizaciones y formatos?

Para resolver esta pregunta, IABIN identificó la necesidad de crear/adoptar estándares para el intercambio de información. En otras palabras que todos los proveedores de datos debían hablar el mismo idioma.

Con esta finalidad, IABIN procedió a identificar o crear estándares que pudieran ser usados o estuvieran siendo usados para el intercambio de datos en cada una de las Redes Temáticas de IABIN y debía crear una estrategia para promover su uso e implementación. El resultado de esta fase se encuentra en el Capítulo 4 “Estándares y Protocolos” de este documento.

- ¿Cómo hacer disponible la información de un proveedor en tiempo real?

Para resolver esta pregunta, se realizó un análisis de las tecnologías existentes, de cómo lo estaban realizando otras redes existentes y se evaluaron cada una de las alternativas encontradas. Al final de este proceso se concluyó que la solución debía de cumplir con las siguientes características:

- o Ser fácil de instalar y mantener. Esto significa que debe permitir conectar los datos de los proveedores de datos de una manera transparente para el proveedor de datos y sin que los proveedores de datos tengan que incurrir en gastos adicionales.

- Garantizar la seguridad de los proveedores de datos y sus datos, y no se debe poner en riesgo la integridad de sus redes internas de información.
- Suministrar la información necesaria para poder identificar el proveedor y sus derechos de propiedad intelectual sobre sus datos.
- Una solución única para toda la red. Esto significa que la solución debería permitir servir datos de Especies, Especímenes, Especies Invasoras, Polinizadores, Áreas Protegidas y Ecosistemas.
- Usar Internet como único canal de comunicación e intercambio de datos. Sin embargo la solución de Internet presenta los siguientes retos:
 - Diferentes velocidades las conexiones entre los proveedores de datos. En teoría, bajo este esquema se contempla que la red siempre va a funcionar a la velocidad del proveedor de datos que posea la conexión más lenta.
 - La disponibilidad de la conexión: Esto significa que el proveedor debería estar disponible cada vez que se realice una consulta, ya que si el proveedor no está en-línea “on-line”, sus datos no podrían ser consultados y recuperados cuando los pida el usuario.
 - No todos los proveedores de datos tendrían una conexión a Internet.
- Permitir que otras redes (si el proveedor así lo desea) a nivel nacional, regional o mundial, puedan tener acceso a sus datos sin necesidad de implementar soluciones paralelas para tal fin y de esta manera evitar la duplicidad de esfuerzos.

Como resultado de lo anterior, se identificó la necesidad de crear/adaptar un “conector”, el que pudiera ser instalado entre el proveedor y la red; y que cumpliera con las características arriba expuestas.

También se identificó que existen dos tipos de datos que la red tiene que integrar y conectar: El primer grupo son los datos primarios sobre biodiversidad (especímenes, especies y observaciones) y el segundo grupo son los datos cartográficos (mapas de los países, mapas de distribución de especies, ecosistemas, etc..)

Para el primer grupo, el Grupo de Trabajo Técnico de IABIN decidió que la mejor alternativa es retomar el trabajo liderando por el grupo de Biodiversity Information Standards (conocido anteriormente como TDWG Taxonomic Database Working Group) y trabajar en conjunto con ellos y GBIF en la implementación de una solución integral para la recuperación de datos llamada TAPIR (TDWG Access Protocol for Information Retrieval). <http://www.tdwg.org/activities/tapir/>, para todos los proveedores de datos de las diferentes redes. El TAPIR es la siguiente generación del proyecto conocido como DiGIR, el cual fue desarrollado con la finalidad de compartir datos biológicos bajo el estándar Darwin Core.

Para el segundo grupo, el Grupo de Trabajo Técnico de IABIN decidió continuar con la estrategia empezada por la Red Geo-Espacial de IABIN, la cual permitía intercambiar información geográfica utilizando estándares como WMS y WFS. Se espera que eventualmente las implementaciones de TAPIR permitan servir los protocolos WMS y WFS

- Cómo saber dónde se encuentran los datos?

Una vez superados los dos pasos anteriores deberíamos de alguna manera crear un mecanismo que nos permitiera realizar las búsquedas en la red y poder recuperar los datos de interés. De igual manera se planteó cual debería de ser la característica de este motor de búsqueda especializado en datos de la red y su diferencia con motores de búsqueda populares tales como Google y que justificara su inversión. Como resultado se creará una herramienta de búsqueda llamada IABIN Biobot (Ver anexo 2. Diferencia entre IABIN Biobot y Google), acompañado de una serie de herramientas y estrategias que serán enmarcadas dentro del Catalogo de IABIN y las cuales permitirán realizar búsquedas en la red. De igual manera se destaca que hay que poner énfasis en la creación de metadatos, para lograr este objetivo.

Esta solución también nos debería permitir crear un índice que facilite la localización de los datos suministrados por cada proveedor. En otras palabras poder determinar con exactitud, que datos existen en la red, que proveedores los poseen y como llegar hasta esto datos.

- Cómo hacer disponible la información que no esta digitalizada?

La red también encontró que gran cantidad de los datos no se encuentran en formato digital ó en documentos electrónicos que no permiten un intercambio automático de datos. Por esta razón la red debería de suministrarles a los proveedores de datos, que quisieran digitalizar sus datos, las herramientas necesarias para tal fin. Bajo este esquema cada Red Temática de IABIN debería adoptar o crear una herramienta, de libre acceso, para la digitalización de datos.

b. Arquitectura Básica de IABIN

Como resultado a las preguntas planteadas originalmente, se obtuvo el esquema básico a ser implementado por cada una de las Redes Temáticas de IABIN (a la medida que sea posible tecnológicamente, según la estructura de la red temática y las características de sus datos) y el cual se muestra en gráfico No. 1.

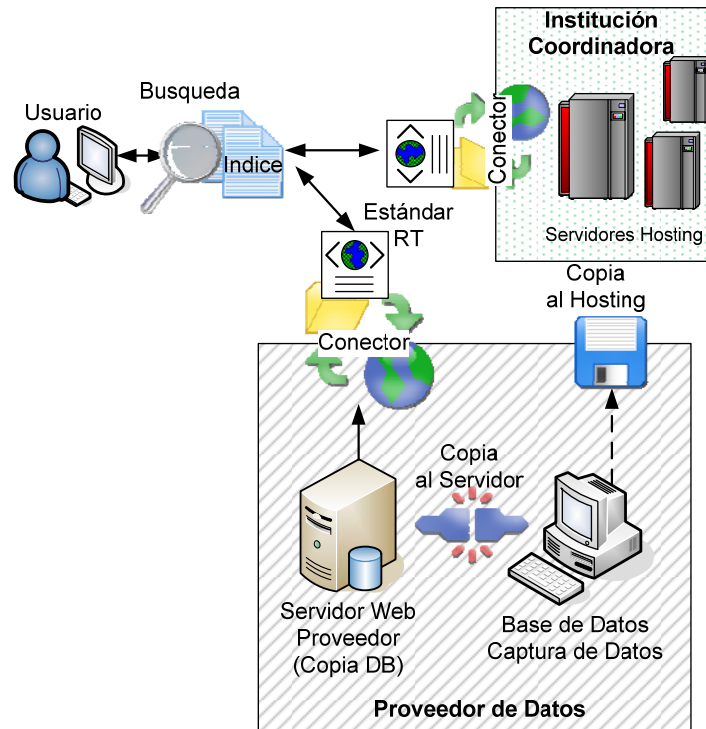


Figura - 1

En este esquema se diferencian los siguientes elementos:

- ✓ **Proveedor de Datos:** Organización dispuesta a compartir sus datos en Internet, siguiendo las políticas y normas establecidas por IABIN. El proveedor de datos continuará teniendo posesión de sus datos. El proveedor de datos podrá compartir sus datos por medio de uno de los siguientes métodos:
 - Conectándose directamente a Internet instalando un “CONNECTOR”. Para esto el proveedor tendrá como mínimo que poseer:
 - Un servidor de Internet
 - Una conexión a Internet con una ancho de banda mínimo de 128kbs
 - O, aprovechando el servicio de hosting que le puede prestar la institución coordinadora. Para esto el proveedor deberá pasar una copia de sus datos a la Institución Coordinadora (o cualquier otra organización, algunas veces conocido como “custodio”, seleccionada por el proveedor y que haga parte de la red y el cual debe de tener un “CONNECTOR” instalado). El Proveedor de datos tendrá la responsabilidad de mantener la información actualizada.

Entre los principales compromisos de los proveedores de datos tenemos:

- Dar acceso a la información
 - Crear y mantener los metadatos de la información que posee
 - Asegurar la calidad de sus datos
 - Mantener actualizada la información
 - Mantener comunicación con el Punto Focal y la Institución Coordinadora
- ✓ **Institución Coordinadora:** Es la institución responsable de la implementación de una de las Redes temáticas de IABIN. Entre sus funciones principales se encuentran:
- Crear un Grupo de Trabajo
 - Definir políticas de uso para la información de la Red Temática
 - Definir, mantener y proveer el estándar(es) de la Red Temática
 - Prestar el servicio de hosting a las organizaciones que así lo requieran y que no tengan capacidad para compartir sus datos en Internet.
 - Capacitación (Cursos y Materiales) sobre la herramienta para digitalizar datos
 - Mantener el portal de Internet la Red Temática
 - Desarrollar la herramienta para digitalizar datos de esa Red Temática
- ✓ **Motor de Búsqueda:** Será el encargado de buscar, localizar e indexar la información suministrada en la red. Esta función va a ser asumida por el Catalogo de IABIN. Este motor de búsqueda también le permitirá a los usuarios buscar información específica sobre una RT (Búsqueda Avanzada)
- ✓ **Conector:** Será el elemento encargado de conectar físicamente al proveedor de datos a la Red:
- Escuchar e interpretar las solicitudes de búsqueda hechas por los usuarios y enviadas por las Redes Temáticas o el Portal de IABIN
 - Buscar la información solicitada en la base de datos del proveedor
 - Devolver la información solicitada en el estándar establecido por la red temática.

Esta arquitectura básica será la misma para la consulta de datos biológicos o geográficos, a los cuales se pueda tener acceso por medio de la red.

3. Implementación de la Arquitectura Básica en cada una de las Redes

En el siguiente capítulo, se describe brevemente como será implementada la arquitectura básica de IABIN, por cada una de las Redes Temática.

a. Red Temática de Especies y Especímenes

Objetivo	Hacer disponible la información existente sobre Especies (Observaciones y distribución de Especies) y Especímenes (Colecciones y observaciones).
Portal de Internet	http://specimens.iabin.net http://especies.iabin.net
Estándares	<ul style="list-style-type: none"> ✓ Especímenes: Darwin Core y ABCD ✓ Especies: Plinian Core
Conector	TAPIR <i>Nota: Proveedores de datos que estén utilizando DÍGIR para suministrar datos de especies, pueden seguir utilizando este conector, si así lo desean.</i>
Digitalización	La red está desarrollando una herramienta propia la cual integrará la información de Especies y Especímenes. <i>Nota: El proveedor de datos es libre de seleccionar la herramienta de digitalización que desee, siempre y cuando este cumpla con los estándares de IABIN.</i>
Integración con otras redes de IABIN	Por medio de: <ul style="list-style-type: none"> ✓ Nombres taxonómicos ✓ Georreferenciación de los especímenes y observaciones. ✓ Mapas de distribución de especies
Retos	<ul style="list-style-type: none"> ✓ Calidad de los datos ✓ Calidad en la Georreferenciación de los datos existes ✓ Una alta cantidad de proveedores de datos ✓ Posible duplicidad en los datos (mismo dato en diferentes proveedores o redes)
Observaciones	✓ Esta red fue el resultado de unir las redes de Especies y Especímenes de IABIN. Esta decisión fue aprobada por el IEC de IABIN, en Junio 2006.

b. Red Temática de Especies Invasoras

Objetivo	Hacer disponible la información existente sobre Especies Invasoras, promoviendo la creación de una base de datos a nivel nacional. Esta red es conocida también como I3N – IABIN Invasive Information Network. Adicional a la información biológica de la especie invasora, esta red está recolectando información sobre impactos económicos y medidas de control de las especies identificadas.
Portal de Internet	http://i3n.iabin.net
Estándares	✓ Estándar I3N <i>Nota: La base del estándar I3N es compatible con el Darwin Core y se le han agregado extensiones para el manejo de datos de interés sólo para las especies invasoras (control, económicos, etc)</i>
Conector	TAPIR (futuras implementaciones) <i>Nota: I3N esta proponiendo implementar solo un cliente de TAPIR para minimizar el impacto en los proveedores de I3N existentes. Las bases de datos nacionales que están funcionando actualmente serán conectadas mas adelante.</i>
Digitalización	La red ha desarrollado su propia herramienta de digitalización. <i>Nota: Los proveedores de datos tienen que utilizar la herramienta de digitalización creada por I3N, ya que esta herramienta captura información de interés para la red (control, economica, etc).</i>
Integración con otras redes de IABIN	Por medio de: ✓ Nombres taxonómicos ✓ Georreferenciación de especímenes y observaciones. ✓ Esta red de comportará como un proveedor de datos de la red de especies y especímenes de IABIN, utilizando los estándares establecidos por esta red.
Retos	✓ Recolección de datos ✓ Calidad en la georreferenciación de los datos existes ✓ Adaptación del software existente para el uso de TAPIR ✓ Poco conocimiento sobre el tema de especies invasoras.
Observaciones	✓ Esta es la red de IABIN que presenta un mayor avance ✓ La estrategia de I3N establece el nombramiento de un líder nacional de I3N y una sola base de datos Nacional, para el manejo y control de especies invasoras en cada país.

c. Red Temática de Polinizadores

Objetivo	Hacer disponible la información existente sobre Polinizadores. Adicional a la información biológica del polinizador, esta red tiene planes futuros de incorporar al sistema la relación existente entre las plantas y sus polinizadores. Esta relación podría tener información económica asociada en un futuro
Portal de Internet	http://pollinators.iabin.net http://polinizadores.iabin.net
Estándares	<ul style="list-style-type: none"> ✓ Especímenes: Darwin Core y ABCD ✓ Especies: Plinian Core. ✓ <i>Nota: La relación entre plantas y sus polinizadores, será manejada como una extensión del estándar Darwin Core. Esta extensión tiene que ser desarrollada y validada.</i>
Conector	TAPIR
Digitalización	Por determinar. <i>Actualmente existen herramientas que permiten la digitalización de colecciones de polinizadores. Se maneja la alternativa de utilizar la misma herramienta que está siendo desarrollada por la Red de Especies y Especímenes de IABIN, y agregar la relación existente entre plantas y polinizadores.</i>
Integración con otras redes de IABIN	Por medio de: <ul style="list-style-type: none"> ✓ Nombres taxonómicos ✓ Georreferenciación de especímenes y observaciones. ✓ Esta red de comportará como un proveedor de datos de la red de especies y especímenes de IABIN, utilizando los estándares establecidos por esta red.
Retos	<ul style="list-style-type: none"> ✓ Recolección de datos ✓ Calidad en la georreferenciación de los datos existes ✓ Pocas bases de datos sobre polinizadores en formato digital ✓ Las herramientas para digitalizar datos de polinizadores, no contienen información propia de la actividad de polinización.
Observaciones	<ul style="list-style-type: none"> ✓ La estrategia primaria de esta red es la de tratar de digitalizar la mayor cantidad de colecciones de polinizadores. ✓ Complejidad de la relación entre plantas y polinizadores.

d. Red Temática de Ecosistemas

Objetivo	<p>Integrar la información existente de los ecosistemas (terrestres, marinos y acuáticos continentales) a nivel regional.</p> <p>Una de las metas principales de esta red es crear un sistema de referencias cruzadas el cual permita realizar equivalencias entre las diferentes clasificaciones utilizadas en el continente.</p> <p>Para lograr este fin se creó un formato estándar (metodología GEOSS) con cinco (5) niveles.</p> <table border="1" data-bbox="516 541 1396 1247"> <thead> <tr> <th>Nivel</th> <th>Tema</th> <th>Terrestre</th> <th>Acuático Continental</th> <th>Marino</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Global Macro-Bioclina & Biogeografía</td> <td>Polar, Boreal, Temperate, Med., Tropical (<i>optional: add vegetation structural response [treed, shrub/herb, semi-desert, tundra, aquatic, sparse]</i>)</td> <td>Polar, Temperate, Tropical</td> <td>Polar, Temperate, Tropical</td> </tr> <tr> <td>2</td> <td>Meso-Bioclina/ Biogeografía</td> <td>Biomes/phytogeography; relevant leaf phenology combinations + phytogeography</td> <td>global biomes + zoogeographic basins/ecoregions (WWF/TNC)</td> <td>marine regions defined by SST, sea surface elevation, mixed layer depth; biogeography</td> </tr> <tr> <td>3</td> <td>Geofísico</td> <td>plant-available soil moisture; hydrogeomorphology, specialized substrates</td> <td>Ecological Drainage Units: climate/physiography/substrate porosity</td> <td>Nearshore FW, Coastal Neritic, Oceanic</td> </tr> <tr> <td>4</td> <td>Estructura Geofísica</td> <td>Vegetation Structure Landscape mosaics</td> <td>Surface water character determining biotic structure (e.g., lakes, rivers, streams)</td> <td>e.g., Estuary, FW plume; Reef; Biomass estimate by water column layer</td> </tr> <tr> <td>5</td> <td>Composición Biótica</td> <td>Gradient from secondary vegetation to undisturbed land cover; Characteristic Genera</td> <td>Macrohabitats; Fish/invert communities</td> <td>Coastal Macrohabitats; Plankton communities</td> </tr> </tbody> </table> <p><i>Nota: Los países seguirán utilizando sus clasificaciones de ecosistemas existentes. El formato estándar es sólo una manera común de describir cada clase.</i></p>	Nivel	Tema	Terrestre	Acuático Continental	Marino	1	Global Macro-Bioclina & Biogeografía	Polar, Boreal, Temperate, Med., Tropical (<i>optional: add vegetation structural response [treed, shrub/herb, semi-desert, tundra, aquatic, sparse]</i>)	Polar, Temperate, Tropical	Polar, Temperate, Tropical	2	Meso-Bioclina/ Biogeografía	Biomes/phytogeography; relevant leaf phenology combinations + phytogeography	global biomes + zoogeographic basins/ecoregions (WWF/TNC)	marine regions defined by SST, sea surface elevation, mixed layer depth; biogeography	3	Geofísico	plant-available soil moisture; hydrogeomorphology, specialized substrates	Ecological Drainage Units: climate/physiography/substrate porosity	Nearshore FW, Coastal Neritic, Oceanic	4	Estructura Geofísica	Vegetation Structure Landscape mosaics	Surface water character determining biotic structure (e.g., lakes, rivers, streams)	e.g., Estuary, FW plume; Reef; Biomass estimate by water column layer	5	Composición Biótica	Gradient from secondary vegetation to undisturbed land cover; Characteristic Genera	Macrohabitats; Fish/invert communities	Coastal Macrohabitats; Plankton communities
Nivel	Tema	Terrestre	Acuático Continental	Marino																											
1	Global Macro-Bioclina & Biogeografía	Polar, Boreal, Temperate, Med., Tropical (<i>optional: add vegetation structural response [treed, shrub/herb, semi-desert, tundra, aquatic, sparse]</i>)	Polar, Temperate, Tropical	Polar, Temperate, Tropical																											
2	Meso-Bioclina/ Biogeografía	Biomes/phytogeography; relevant leaf phenology combinations + phytogeography	global biomes + zoogeographic basins/ecoregions (WWF/TNC)	marine regions defined by SST, sea surface elevation, mixed layer depth; biogeography																											
3	Geofísico	plant-available soil moisture; hydrogeomorphology, specialized substrates	Ecological Drainage Units: climate/physiography/substrate porosity	Nearshore FW, Coastal Neritic, Oceanic																											
4	Estructura Geofísica	Vegetation Structure Landscape mosaics	Surface water character determining biotic structure (e.g., lakes, rivers, streams)	e.g., Estuary, FW plume; Reef; Biomass estimate by water column layer																											
5	Composición Biótica	Gradient from secondary vegetation to undisturbed land cover; Characteristic Genera	Macrohabitats; Fish/invert communities	Coastal Macrohabitats; Plankton communities																											
Portal de Internet	<p>http://ecosystems.iabin.net http://ecosistemas.iabin.net</p>																														
Estándares	<p>✓ Formato estándar para la descripción de un ecosistema.</p>																														
Conector	<p>✓ WS (Servicios de Web) para el formato estándar y la referencia cruzada.</p> <p>✓ WFS para el acceso a datos geográficos.</p>																														
Digitalización	<p>Esta red creó su propia herramienta para la digitalización de datos. Esta herramienta permite llenar el formato estándar de la red.</p>																														
Integración con otras redes de IABIN	<p>Por medio de:</p> <p>✓ Coordinadas geográficas</p> <p>✓ Listado de especies dominantes en el ecosistema</p> <p>✓ Integración geo-espacial.</p>																														
Retos	<p>✓ Recolección de datos</p>																														

	<ul style="list-style-type: none">✓ La creación de las referencias cruzadas.✓ Varios tipos de clasificaciones de ecosistemas son utilizados en el continente, lo cual imposibilita una equivalencia de un 100% entre un ecosistema y otro.✓ Dificultad en llenar la información de Nivel 5, del formato✓ Gran cantidad de la información referente a Ecosistemas se encuentra en mapas.✓ Existe información sobre ecosistemas terrestres, pero la información de los ecosistemas marinos y acuáticos continentales es muy escasa.
Observaciones	<ul style="list-style-type: none">✓ Se espera que las redes de Especies y Especímenes, Especies Invasoras y Polinizadores, proporcionen y digitalicen la información necesaria (ocurrencias con coordenadas o límites) para poder determinar las especies existentes dentro de un ecosistema.✓ Esta red no digitalizará datos de especies y especímenes existentes en el ecosistema.

e. Red Temática de Áreas Protegidas

Objetivo	Hacer disponible la información referente sobre áreas protegidas, teniendo como prioridad la información referente a el manejo de estas áreas.
Portal de Internet	http://protectedareas.iabin.net http://areasprotegidas.iabin.net
Estándares	<i>Por aprobar</i> ✓ WDPA Core Ver. 1.2 (World Data Base Protected Areas)
Conector	<i>Por aprobar</i> ✓ TAPIR. ✓ WFS para el acceso a datos geográficos.
Digitalización	Por determinar.
Integración con otras redes de IABIN	Por medio de: ✓ Coordinadas geográficas ✓ Listado de especies dominantes en el área protegida. ✓ Integración geo-espacial.
Retos	✓ Recolección de datos y actualización de la nueva versión del WDPA Core. (Versión 1.2) ✓ Gran cantidad de la información sobre Áreas Protegidas se encuentra en mapas.
Observaciones	✓ Se espera que las redes de Especies y Especímenes, Especies Invasoras y Polinizadores, proporcionen y digitalicen la información necesaria (ocurrencias con coordenadas o límites) para poder determinar las especies existentes dentro de un Área Protegida. ✓ Esta red no digitalizara datos de especies y especímenes existentes en el área protegida.

f. Red Geo-Espacial

Objetivo	Hacer disponible la información cartográfica existente. <i>Nota: Esta red no existía originalmente en IABIN. Se creo viendo la necesidad de acceso a la información geográfica existente, tal como: Limitse de países, ciudades, ríos, lagos, etc.</i>
Portal de Internet	http://geoespacial.iabin.net http://geoespacial.iabin.net
Estándares	✓ FDGC (Estandar de metadatos espaciales)
Conector	✓ WFS para el acceso a datos geográficos.
Digitalización	No aplica.
Integración con otras redes de IABIN	Por medio de: ✓ Coordinadas geográficas ✓ Integración geo-espacial.
Retos	<ul style="list-style-type: none"> ✓ Estandarización en la presentación de los diferentes mapas ✓ En la integración de mapas de diferentes países es necesario llegar a acuerdos sobre fronteras ✓ No existen una institución coordinadora responsable de esta red Se espera que este rol sea liderado por las redes de Ecosistemas y Áreas Protegidas ✓ Dificultad para tener acceso a la información cartográfica oficial de cada país ✓ Necesita de altas velocidades en las conexiones a Internet
Observaciones	<ul style="list-style-type: none"> ✓ Su implementación se realiza con la instalación e integración de servidores de mapas nacionales. ✓ Es posible que esta red desaparezca en el futuro y las redes de Ecosistemas y Áreas Protegidas asuman un rol compartido de mantener esta red.

g. Catálogo

Objetivo	<p>Integrar y facilitar la búsqueda de la información suministrada por la Red. El Catálogo de IABIN proporcionará los siguientes servicios:</p> <ul style="list-style-type: none"> ✓ IABIN Biobot: Motor de búsqueda para la recuperación de datos biológicos, en tres idiomas (inglés, español y portugués) ✓ UDDI: sitio de Registro de proveedores de IABIN ✓ Índice Geográfico (Gacetero) ✓ Índice Organizacional ✓ Frases Comunes ✓ Tesauros ✓ Registro de Metadatos ✓ Registro de Proveedores de datos Espaciales. <p><i>El Catálogo tendrá la capacidad de leer e integrar bases de datos de metadatos existentes en los países.</i></p> <p><i>El motor de búsqueda realizara la búsqueda de una palabra en inglés, español y portugués, gracias al tesoro y otras listas controladas multi-lenguaje, pero el contenido será mostrado en su idioma original (No será traducido)</i></p>
Portal de Internet	No Aplica.
Estándares	<ul style="list-style-type: none"> ✓ FDGC (Estandar de metadatos espaciales) ✓ Dublin Core (Estandar para documentos, imágenes)
Conector	✓ Servicios de Web (WebServices)
Digitalización	✓ Para la digitalización de metadatos.
Integración con otras redes de IABIN	<p>El catálogo de IABIN podrá realizar búsquedas por medio de:</p> <ul style="list-style-type: none"> ✓ Nombres taxonómicos ✓ Nombres comunes ✓ Frases ✓ Coordenadas geográficas y Nombres Geográficos
Retos	<ul style="list-style-type: none"> ✓ Necesita de altas velocidades en las conexiones a Internet. ✓ Calidad en los metadatos ✓ Poco desarrollo en la creación de metadatos.
Observaciones	<ul style="list-style-type: none"> ✓ IABIN creará un tesoro centralizado el cual se alimentara de lo tesauros regionales. Cada término será introducido en tres idiomas. ✓ El tesoro esta disponible también a través de servicios web y para uso/adopción de todas las RTs y de la comunidad los cuales puede ser incluida en sus aplicaciones y sistemas

4. Integración de las Redes

Hasta el momento hemos descrito cómo IABIN va a integrar la información del mismo tipo (ej. Especies invasoras) que poseen los diferentes proveedores de datos.

Pero con el fin de que la información suministrada por cada una de las Redes Temáticas de IABIN sea realmente útil para la toma de decisiones, se hace necesario que los diferentes tipos de datos puedan ser integrados.

La capacidad de desarrollar herramientas que integren de forma homogénea esta variedad de datos nos permitirán tomar mejores decisiones. Por tanto uno de los grandes retos de IABIN es el desarrollo de estas herramientas.

También es importante que todos los socios de IABIN entiendan la importancia de los estándares para la digitalización de datos, ya que estos estándares son los que permitirán esta integración

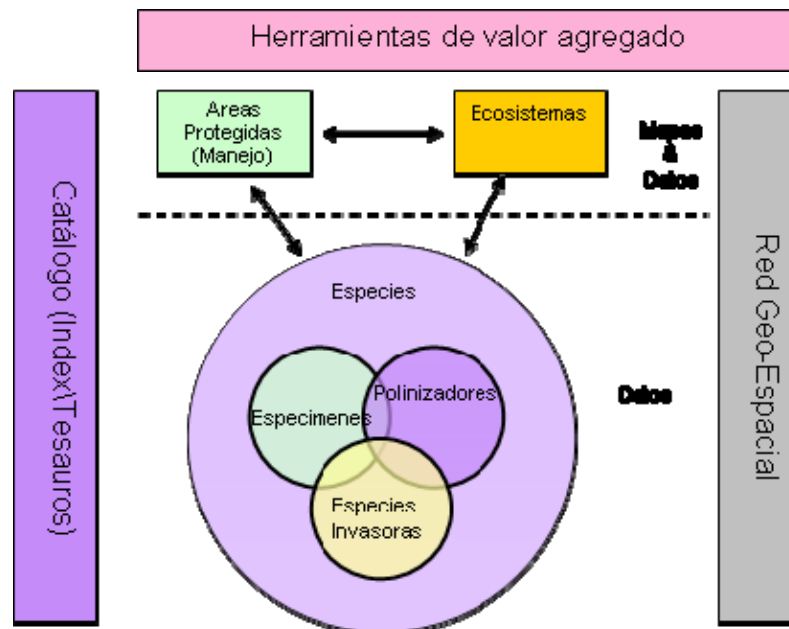


Figura - 2

a. Concepto Básico de Integración.

Los datos que serán suministrados por los diferentes proveedores de datos de IABIN son básicamente de dos tipos: Los datos biológicos primarios y datos espaciales. Estos tipos de datos pueden integrarse bajo las siguientes combinaciones:

- ✓ Datos biológicos con datos biológicos: Este esquema es el que va a permitir integrar información existente de las diferentes especies y especímenes. Para que este esquema pueda ser aplicado se hace necesario utilizar descriptores comunes para los datos.

- ✓ Datos biológicos con datos espaciales: Este esquema es el va a permitir integrar la información de especies y especímenes, bajo el contexto de un mapa. Para que este esquema funcione se hace necesario que la información biológica tenga una representación en coordenadas que permitan localizar este dato en el espacio. A esta localización de un dato en el espacio se le llama georreferenciación.
- ✓ Datos espaciales con datos espaciales: Este esquema es el que va a permitir integrar la información contenida en dos mapas, generando un nuevo mapa.

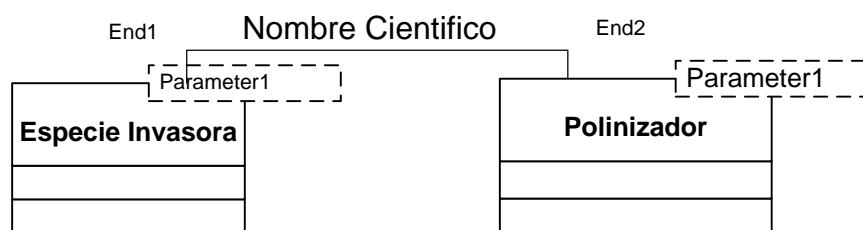
Como resultado de todo el proceso descrito anteriormente tenemos la Arquitectura de IABIN, la cual se muestra en el gráfico No. 3. Este gráfico muestra la forma comocada una de las Redes Temáticas y proveedores de datos van a exponer sus datos para que las demás redes las puedan consultar.

Debido a la gran variedad de las posibles combinaciones que se pueden generar al cruzar los diferentes tipos de datos, el Grupo de Trabajo Técnico de IABIN ha sugerido que esta primera fase la Red se enfoque en solucionar las integraciones básicas para cada uno de los uno de los esquemas de integración. Estas integraciones serán:

1 – La integración de datos biológicos con datos biológicos. Solo se realizará utilizando los nombres científicos de las especies y especímenes que estén en las diferentes bases de datos. La razón por la cual sólo se va a utilizar el nombre científico es que éste es el único elemento de todo el registro que de alguna manera ha sido estandarizado a nivel internacional y se están creando bases de datos de autoridades taxonómicas las cuales permitirán validar la información existente, además, de que el nombre científico de una especie no se ve afectado por el idioma.

Es bueno aclarar que el Catálogo de IABIN podrá realizar búsquedas en todo el registro, como por ejemplo utilizando el nombre común, pero la integración solo tendrá en cuenta los nombres científicos para determinar si el registro hace referencia a la misma especie o no.

La razón por la cual no se utilizarán otros elementos tales como nombre común, es que estos pueden variar de una región a otra o de un país a otro, además de que idiomáticamente es mucho más complejo realizar búsquedas.



2 – *La integración de datos biológicos con datos espaciales.* Solo se realizará con aquellos registros biológicos que estén georreferenciados, posean coordenadas geográficas, que permitan identificar el lugar en el cual fue colectada u observada la especie.



3 – *La integración de datos espaciales con datos espaciales.* Solo se realizará por medio de la superposición de mapas.

Arquitectura de IABIN

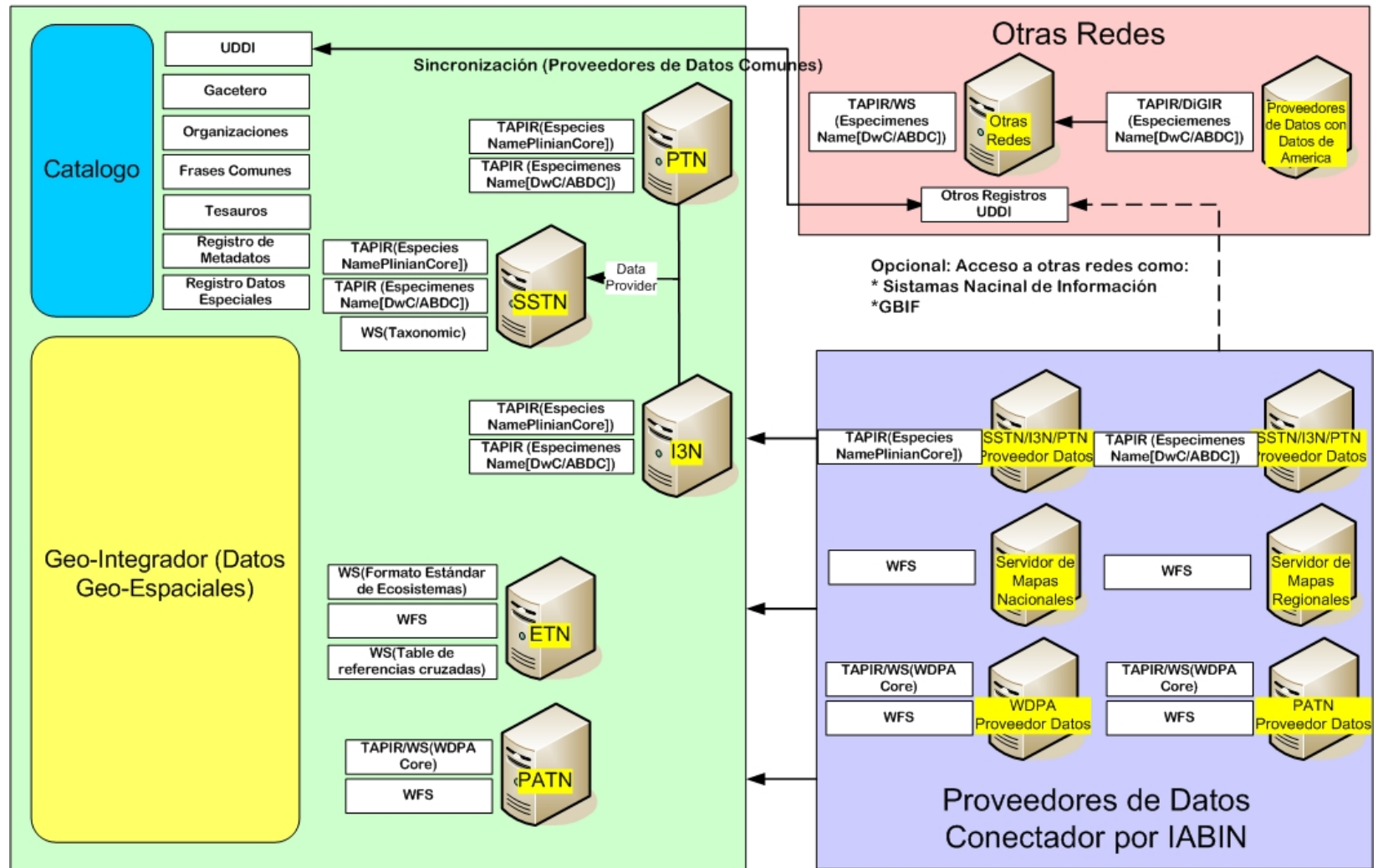


Figura - 3

b. Casos de uso en la Primera Fase (Proyecto IABN-GEF)

En la primera fase de implementación de IABIN, se podrán realizar cuatro tipos de consultas integradas. Estas consultas son llamadas Casos de Uso, los cuales serán:

1 – *Buscar una palabra o frase utilizando el IABIN Biobot.* Se espera que esta búsqueda sea la más común. En esta búsqueda el usuario digitará una palabra tal como “especies invasoras” y el IABIN Biobot recuperará toda la información existente en la red relacionada a esta frase en inglés, español y portugués. La información recuperada puede ser de cualquier tipo: bases de datos, datos tabulares, documentos, imágenes, etc..

2 - *Búsqueda de especies y especímenes utilizando el nombre científico.* El usuario digitará el nombre científico de una especie y el motor de búsqueda desarrollado por la Red de Especies y especímenes, recuperará todos los registros disponibles en la red sobre esa especie.

3 – *Buscar la información existente dentro de un área geográfica seleccionada por el usuario.*

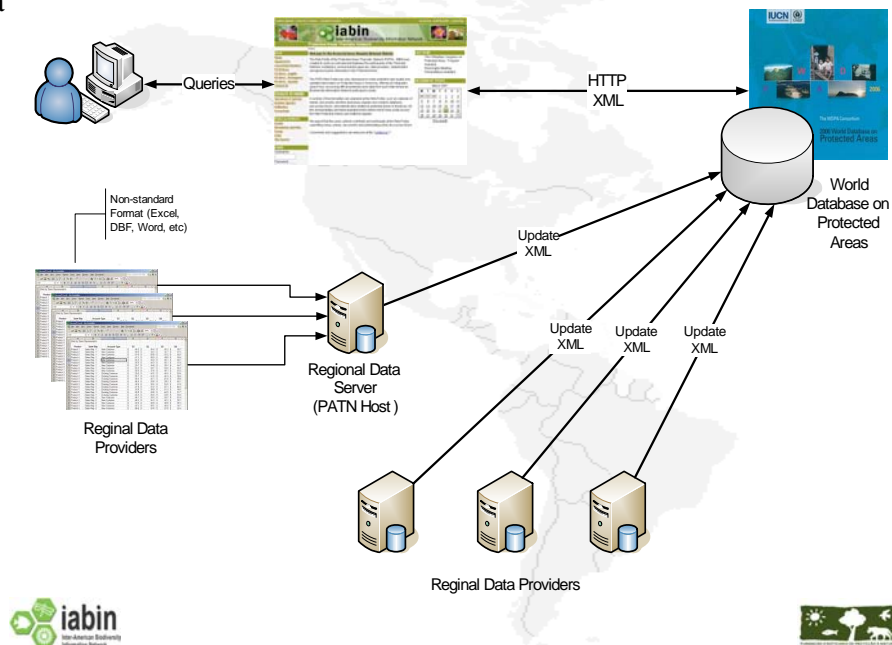
4 – *Buscar la información relacionada, en “X” kms a la redonda, a un punto geográfico determinado por el usuario.*

c. Integración con Otra Iniciativas

Se prevé que la información suministrada por IABIN, podrá ser integrada a otras iniciativas, siempre y cuando éstas utilicen los mismos estándares y protocolos que IABIN esta utilizando para el intercambio de datos.

Cuando un proveedor de datos es conectado a la Red, cualquier otra red lo podrá consultar, siempre y cuando el proveedor de datos esté de acuerdo.

Como un ejemplo de esto es la arquitectura propuesta por PATN en esta integrada con la Base de Da



5. Estándares y Protocolos

En la cuarta reunión de consejo de IABIN, en abril de 2005, los protocolos y estándares iniciales de IABIN fueron aprobados.

Dado que una tarea continua del Grupo de Trabajo Técnico de IABIN es revisar los estándares aprobados y actualizarlos. La siguiente es la tabla de los estándares y protocolos de IABIN actualizada, según la última reunión en marzo de 2007

Parte de la Arquitectura de IABIN	Estándar o Protocolo Adoptado
Arquitectura	Web Services
Servicios de Registro	UDDI
Descripción de Interfase	WSDL
Protocolos de Acceso	TAPIR DiGIR (Si el proveedor lo tiene instalado)
Codificación de Datos	XML
Transporte	HTTP over TCP/IP
Metadatos	
o Para Datos Bibliográficos	Dublin Core
o Para Colecciones y Observaciones de Especímenes	Darwin Core ABCD Schema
o Para Especies	Plinian Core
o Para Áreas Protegidas	WDPA Core Versión 2
o Para Especies Invasoras	Estándar I3N
o Para Datos Espaciales	FGDC
o Para Recursos Biológicos Generales	CSDGM with Bio Profile
o Para procesamiento de Datos Geográficos	Open GIS Consortium (OGC) WFS WMS (Solo si WFS no está disponible)
Para formato de documentos	HTML, PDF, y ASCII
Formatos Gráficos	PNG, JPEG, GIF, WebCGM

6. Consideraciones Generales de la Arquitectura de IABIN.

La arquitectura presentada en este documento y en esta Reunión del Consejo de IABIN, es una arquitectura que se encuentra en continua evolución y se debe ir ajustando según los requerimientos tecnológicos de cada Red Temática y de nuevas tecnologías.

Sin embargo, hay algunas consideraciones generales:

a. Calidad de datos

Garantizar la calidad de los datos conectados a la red, debe de ser una de las metas principales de IABIN. Sólo de esta manera se podrá garantizar que los datos puedan ser utilizados para la toma de decisiones.

Para IABIN es imposible garantizar cada dato que es colocado en la Red, pero cada Red Temática va ha establecer mecanismos de retroalimentación por medio

de los cuales se les informará a los proveedores de datos sobre los errores encontrados en sus datos para que éstos puedan ser corregidos.

Este mecanismo permitirá que la herramienta de búsqueda de IABIN, el IABIN Biobot, le de mayor prioridad (mayor peso) a aquellos proveedores de datos catalogados como altamente confiables y de igual manera bajar en prioridad a aquellos proveedores de datos que sean reportados en ocasiones sucesivas con errores en sus datos.

Con este mecanismo la responsabilidad de la calidad de los datos, recaerá directamente sobre los mismos proveedores de datos.

IABIN, por medio de su componente de capacitación se enfocará en capacitar a los proveedores de datos en como corregir sus datos. Los problemas más comunes encontrados en los datos son de dos tipos:

- i.** Los datos biológicos no cumplen con los estándares establecidos. Esto se debe a que la mayoría de los datos fueron digitalizados cuando estos estándares no existían, por tal motivo los datos tienen que ser completados según el estándar.
- ii.** Georreferenciación de los datos existentes. Las metodologías de georreferenciación utilizando mapas digitales o sistemas de posicionamientos globales (GPS) son relativamente nuevas. Por esta razón, gran parte de los datos colectados con anterioridad poseen problemas de este tipo los cuales tienen que ser corregidos, para poder ser utilizados en procesos de toma de decisiones.

7. Anexos

Anexo 1: Principios de IABIN

IABIN ha adoptado 11 principios rectores para la interoperabilidad de formatos, estándares y protocolos

1. Acceso irrestricto a todos los tipos de datos e información de IABIN sin importar donde resida e ínter-operable con CBD-CHM, GBIF y otras redes;
2. Estándares abiertos, de amplio soporte y no-propietarios;
3. Compatibilidad con estándares emergentes de la principales redes regionales, globales y nacionales de información biológica;
4. Minimización de las restricciones tecnológicas impuestas por las arquitecturas de redes;
5. Desarrollo incremental por fases;
6. Escalable; de modo que permita que los estándares puedan ser usados y aplicados a diversas escalas de redes: global, regional y nacional;
7. Inclusión (ej: facilitar búsquedas en el idioma local) in el diseño de las aplicaciones;
8. Experiencia y capacidad de intercambio a por medio de la red;
9. Respeto por los Derechos de Propiedad Intelectual, Conocimiento Tradicional, Derechos y Reglas para el acceso y beneficio de intercambio de Fuentes Genéticas de acuerdo con el principios y directrices del CBD, y legislaciones nacionales;
10. Extensibilidad y compatibilidad futura y hacia atrás;
11. Minimización de los costos, asegurando un servicio confiable a los usuarios.

Anexo 2. Diferencia entre IABIN Biobot y Google

(Autor: Mike Frame, USGS)

No todas las herramientas de búsqueda son iguales debido a varias características, incluyendo el contenido/alcance de la herramienta de búsqueda, los algoritmos para establecer jerarquías que sean usados, o el uso de vocabularios controlados o de terminologías que apoyen la recuperación de datos o información. Abajo presentamos una comparación básica entre el IABIN BioBot y una herramienta de búsqueda típica tal como Google.

Google:

- La misión de Google es hacer visible todo el universo de la Web y poder encontrar la información contenida en el mismo. Esto no incluye contenidos que no tienen como base la Web que se encuentran en bases de datos o servidores de dominio científico restringidos.
- Google no puede interpretar el contexto de la búsqueda de un usuario. Google lleva a cabo una búsqueda y no conoce el dominio para la misma. Por ejemplo, un pedido por el término “halcones” en Google devolvería resultados para equipos deportivos, especies y/o una estrella de cine.
- Google no entiende el contenido de un documento ni cual es su intención o contexto general.
- El algoritmo para establecer jerarquías de Google está basado principalmente en vínculos externos; recursos vinculados mayormente a otros similares reciben una jerarquía más alta. Este método para recuperar datos científicos puede no ser el mejor método para asegurar que ocurre duplicación de investigación o investigación innovadora.

IABIN BioBot:

- La misión de IABIN es hacer más fácil de encontrar la Web biológica o ecológica.
- El IABIN BioBot interpreta el contexto dentro de los metadatos (sujeto, taxonomía, geografía) para ayudar a presentar un acceso más correcto y distribuido a los recursos.
- El IABIN BioBot entiende a qué se refiere un documento debido a que su dominio y/o archivos de datos primarios se relacionan a las ciencias biológicas.
- El algoritmo para establecer jerarquías de IABIN BioBot está basado en la asignación de diferentes pesos a los campos de metadatos y asigna mayor peso a fuentes que dicen “Confiable/Alta Calidad/Aceptable” que a los contenidos “no asociado”. Esto ayuda a asegurar una mayor calidad y resultados relevantes específicos a las preguntas del usuario.
- Los resultados de IABIN BioBot son presentados en una interfase con Tabs que permite al usuario ver “Todos” los resultados por norma, o seleccionar tipos específicos de recursos (por ejemplo Mapas, Imágenes, Artículos de Revistas Científicas, etc.). esto ayuda a categorizar la información a los usuarios en vez de simplemente presentar una larga lista narrativa de los resultados.

- El IABIN BioBot utiliza un tesoro siempre presente para ayudar a los usuarios a recuperar información que puede no haber sido catalogada o descrita en el mismo contexto o con la misma terminología usada en la pregunta del usuario. Por ejemplo, si un usuario pregunta por “especies invasoras”, las herramientas de búsqueda típicas de la Web solo pueden recuperar datos/información que contenga esa frase exactamente. A través del uso de tesauros, los usuarios del BioBot obtendrán resultados de términos relacionados tales como “especies no nativas”, “especies alienígenas”, etc.