



CONABIO

GOBIERNO  
FEDERAL

Comisión Nacional  
para el Conocimiento y  
Uso de la Biodiversidad



SNIB

Sistema Nacional de Información sobre Biodiversidad



# Control de calidad en El Sistema Nacional de Información Sobre Biodiversidad (SNIB)

Liliana Lara Morales

11 de octubre 2012



# Información vs. Sabiduría

¿Dónde está la sabiduría?

*Perdida en el conocimiento*

¿Dónde está el conocimiento?

*Perdido en la información*

*T. S. Eliot*

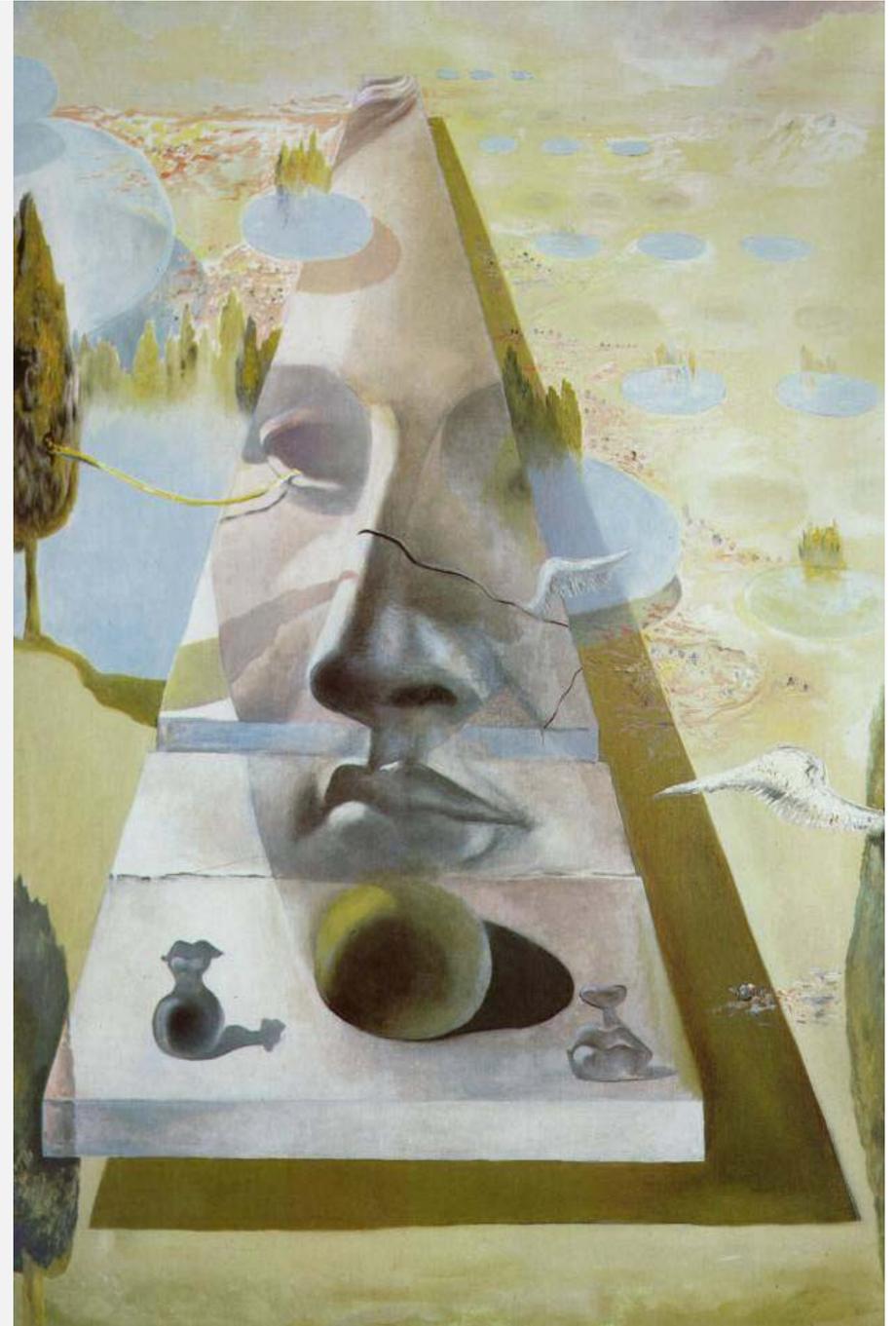
¿Dónde está la información?

*Perdida en los datos*

¿Dónde están los datos?

*Perdidos en la @#~& base de datos*

*Joe Celko*



# Sistema de información



Conjunto de elementos ordenados y relacionados entre sí que contribuyen al desarrollo de un determinado objetivo

El conocimiento tradicional lo constituyen saberes que pertenecen a un pueblo o comunidad indígena; en general se caracteriza por ser una creación colectiva, expresada en una lengua particular (Toledo, 2001).



# Cómo se construye el SNIB

La CONABIO ha construido el SNIB, principalmente con el apoyo a proyectos y con la repatriación de datos, convocando a la comunidad científica relacionada con el estudio de la biodiversidad

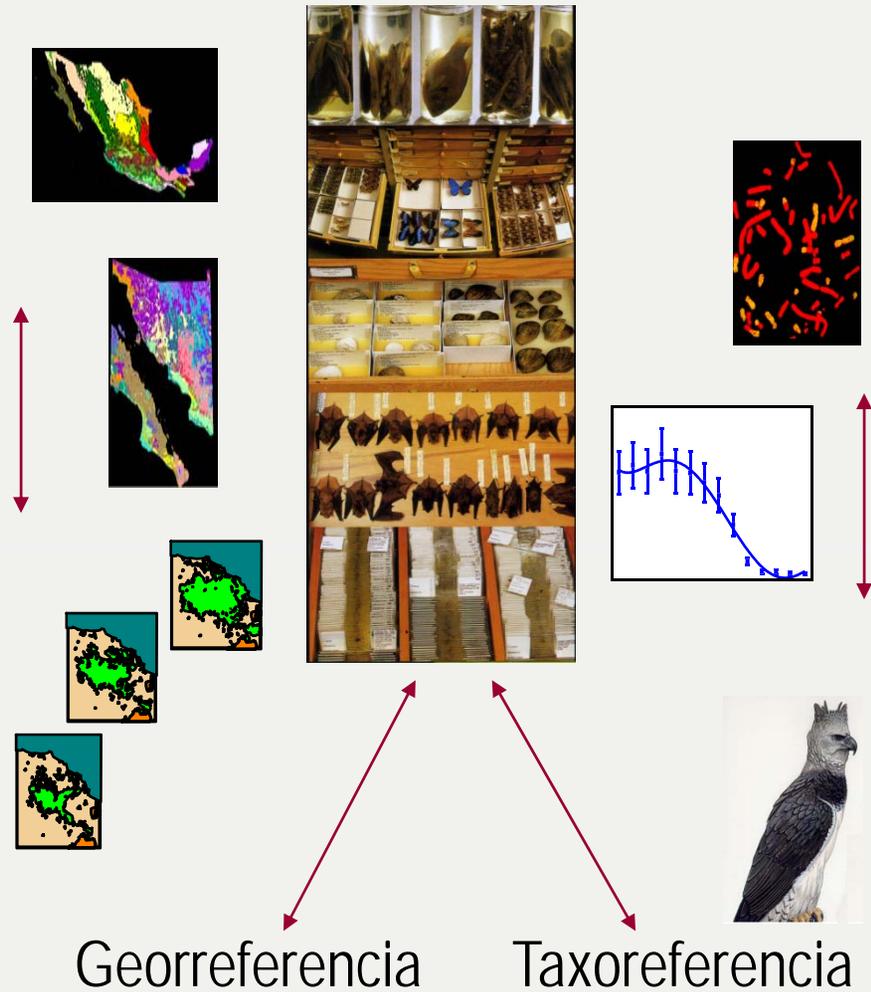
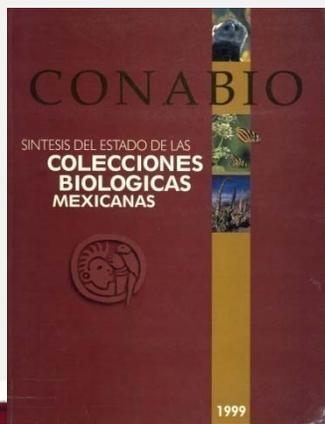
## Tipo de proyectos que se apoyan:

- Inventarios ( florísticos, faunísticos, en México, en áreas naturales protegidas, regiones terrestres, marinas o hidrológicas prioritarias, AICAS, etc.)
- Computarización de colecciones, repatriación de datos
- Fichas de especies (NOM o CITES, especies invasoras, especies prioritarias)
- Catálogos de autoridad taxonómica (con o sin ejemplares respaldo)



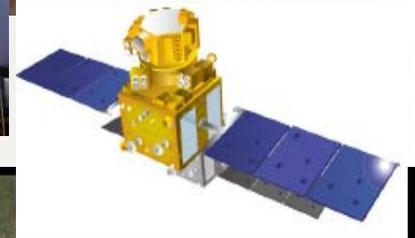
# Cómo se construye el SNIB

El primer eje del SNIB está constituido con base en las especies representadas por los ejemplares depositados en colecciones biológicas y por ejemplares observados y avalados por un especialista



# Cómo se construye el SNIB

El segundo eje de información del SNIB son los sistemas de información geográfica (cartografía digital y la percepción remota). Son herramientas indispensables para el análisis espacial de los patrones de biodiversidad



## Características del Sistema Nacional de Información sobre Biodiversidad

Accesibilidad	Datos a disposición del público, con mecanismos de acceso claros. Cualquier persona pueda realizar sus análisis
Transparencia	Formatos, contenidos (diccionario de datos) citas de las fuentes de los datos. Dar crédito a los autores y poder consultar la fuente original.
Interoperabilidad	Datos compartidos con sistemas informáticos o bancos de datos mediante la utilización de estándares internacionales (REMIB, GBIF, MOBOT, ITIS).

## Características del Sistema Nacional de Información sobre Biodiversidad

### Relevancia

Los datos integrados deben ser relevantes para construir el conocimiento adecuado y poder responder a las preguntas que se planteen. La relevancia de los datos es un tema dinámico que enriquece al sistema.

### Confiability

**Los protocolos de control de calidad establecidos para el sistema garantizan la calidad de los datos. Es un hecho que no existen sistemas sin errores, pero es posible minimizarlos y caracterizarlos.**

### Verificabilidad

La generación de los datos y de la información debe poseer rigor de la investigación científica. Las opiniones son válidas, pero deben estar sustentadas en datos factuales para incorporarlos al sistema



CONABIO

GOBIERNO  
FEDERAL

## Estándares

Asegurar la calidad e interoperabilidad de datos de diversas fuentes

### SNIB. Instructivo conformación de bases de datos, obligatoriedad.

***Datos curatoriales.*** Recolector, no. de colecta, fechas de recolecta u observación y determinación, número de individuos o duplicados, tipo de preparación ambiente, sitio de recolecta u observación (latitud-longitud) método de obtención de la coordenada, datum, descripción de la localidad, país, estado, municipio, tipo de vegetación, determinación, número de catálogo, colección de referencia.

***Datos nomenclaturales.*** Nombre científico, autoridad año de publicación, ubicación dentro de una jerarquía taxonómica, sistema de clasificación o catálogo de autoridad, estatus del taxón.



# Estándares

El SNIB es compatible con Darwin Core

## SNIB. Instructivo conformación de bases de datos, obligatoriedad.

**Datos de especies.** Información biológica, distribución, ecología.

**Referencias bibliográficas.** Autor(es), año, título, editorial, volumen, número, páginas.

**Restricción de la información.** Tipo de dato restringido (todo el registro, localidad, coordenada, región, imagen, ninguna, toda) fecha y motivos de la restricción.

## Casos particulares.

**Datos etnobiológicos.** Nombre común o vernáculo, lengua, usos.

**Regionalización.** Tipo y nombre de región (ANP, región terrestre prioritaria, región hidrológica prioritaria, ecorregión).

**Objetos externos.** Fotografías digitales, ilustraciones científicas, paisaje, archivos pdf., doc., mapas digitales.



# Estándares

- **Nomenclatura biológica**

Códigos de nomenclatura (plantas CINB, animales CINZ, bacterias CINB, plantas cultivadas CINCP).

- **Autores de nombres de plantas**

Brummitt, R. K., & C. E. Powell (Eds.). 1992. Authors of plant names. Royal Botanic Gardens, Kew, 732 pp.

Pichi Sermolli, R. E. G. 1996. Authors of scientific names in Pteridophyta. Royal Botanic Gardens, Kew, 78 pp.

Villaseñor Ríos, J. L., E. Ortiz & R. Redonda-Martínez. 2009. Catálogo de autores de plantas vasculares de México. Universidad Nacional Autónoma de México, Instituto de Biología, Conabio. México. 69 pp



# Estándares

- **Información Geográfica**

México. Instituto Nacional de Estadística y Geografía (INEGI). División política nacional (País/Estado/Municipio) Localidades.

Validación geográfica. Sobreposición cartográfica.

- **Georreferencia**

Comisión Nacional para el Conocimiento y uso de la Biodiversidad (CONABIO). 2006. Manual de procedimientos para Georreferenciar localidades de colecciones Biológicas. México. 169 págs.



# Estándares

- **Sistemas de clasificación de tipos de vegetación de México**

Conabio-Comité Asesor del Proceso de Montreal. Obtenido de: Instituto Nacional de Estadística Geografía e Informática. 2002. Conjunto de datos vectoriales de la carta de uso de Suelo y vegetación Serie II. Escala 1:250,000. Conjunto nacional. México.

Flores, G., J. Jiménez, X. Madrigal, F. Moncayo, F. y F. Takaki. 1971. Mapa y descripción de los tipos de vegetación de la República Mexicana. Secretaría de Recursos Hidráulicos, México.

INEGI 2002. Carta de vegetación primaria de México escala 1:1,000,000.

Miranda, F. y E. Hernández, X. 1963. Los tipos de vegetación de México y su clasificación. Boletín de la Sociedad Botánica de México 28: 29-179.

Rzedowski, J. 1978. Vegetación de México. Limusa. México. 432 pp.



# Control de calidad

Desarrollo de rutinas y procedimientos para el control de calidad de los datos. Es muy poco apreciado y merece ser subrayado que los datos obtenidos de los museos y herbarios necesitan ser evaluados para detectar inconsistencias y errores (Soberón *et al.*, 2001)

## 4.2 Ejemplares repetidos. *Eliminar los registros repetidos o aclarar.*

Id Ejemplar	Categoría Taxonomica	Nombre	Siglas Colección	No De Catalogo	No De Colecta	Fecha De Colecta	Id Grupo	Grupo Colectores/ Observadores
1959	especie	Abeilliaabeillei	IHNORNI	1230	36	23/ 11 / 1999	2	Hunn E.
1961	especie	Abeillia abeillei	IHNORNI	1230	36	23/ 11 / 1999	2	Hunn E.
1963	especie	Abeillia abeillei	IHNORNI	1230	36	23/ 11 / 1999	2	Hunn E.
24755	especie	Acacia angustissima	CHAPA	12345	134	11 /11/ 1945	2	Bachem C. U. & Rojas C. R.
24756	especie	Acacia angustissima	CHAPA	12345	134	11 /11/ 1945	2	Bachem C. U. & Rojas C. R.
30907	especie	Adiantum capillus-veneris	MEXU	23456	1237	07 / 10 / 2008	2	Tejero
31207	especie	Adiantum capillus-veneris	MEXU	23456	1237	07 / 10 / 2008	2	Tejero

# ¿quién hace el control de calidad?

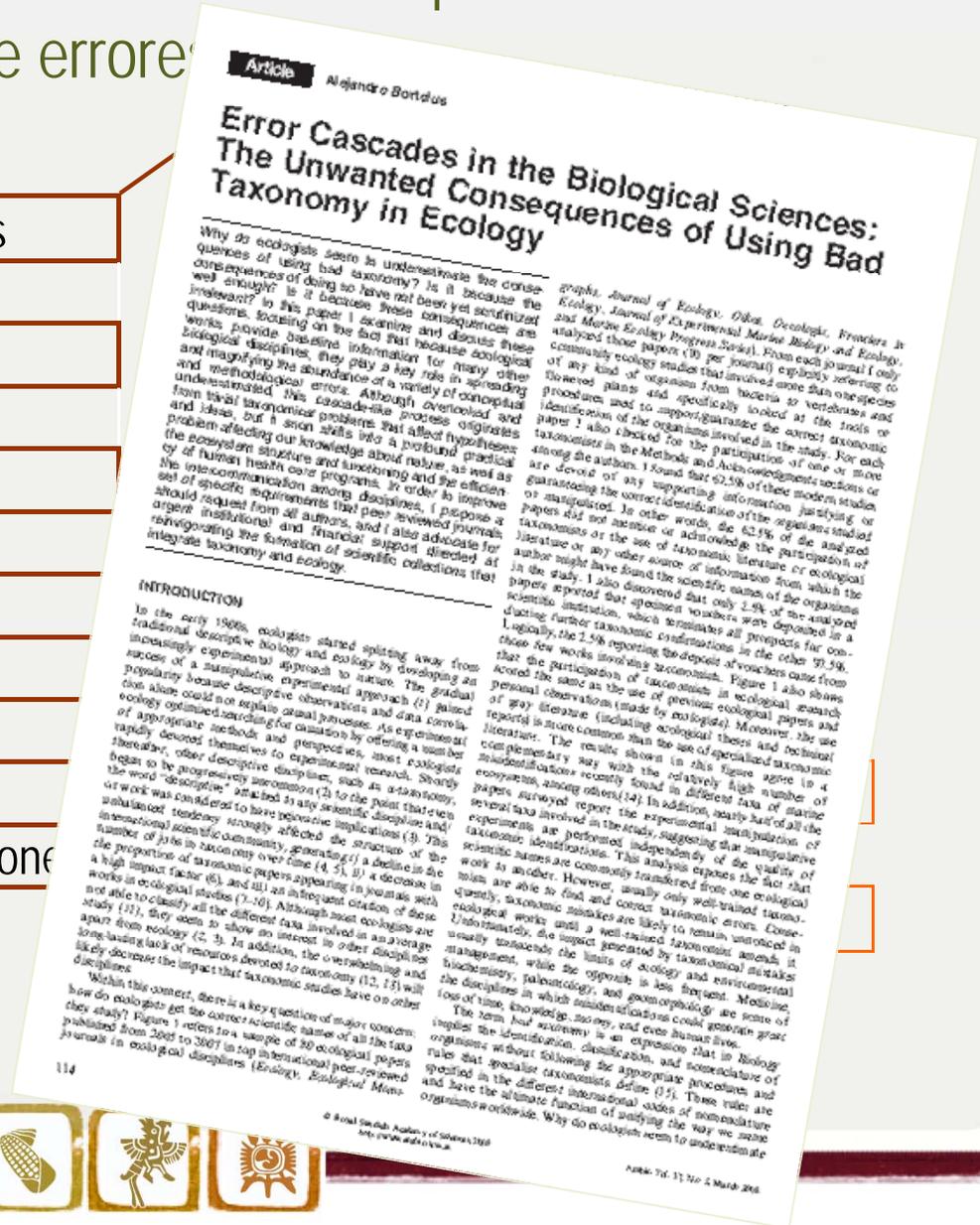
En la CONABIO Un grupo de 12 biólogos especializados, realizan la evaluación de los contenidos en las bases de datos



El especialista encargado del proyecto financiado por la CONABIO, instruye a sus técnicos y capturistas, para el cotejo de los los posibles errores y su corrección en la base de datos

Para realizar el control de calidad en cada capa de información se han distinguido siete tipos de errores

Capa de información (6)	Personas y grupos
	Ejemplar
	Taxonómica
	Geográfica
	Bibliográfica
	Colecciones e instituciones



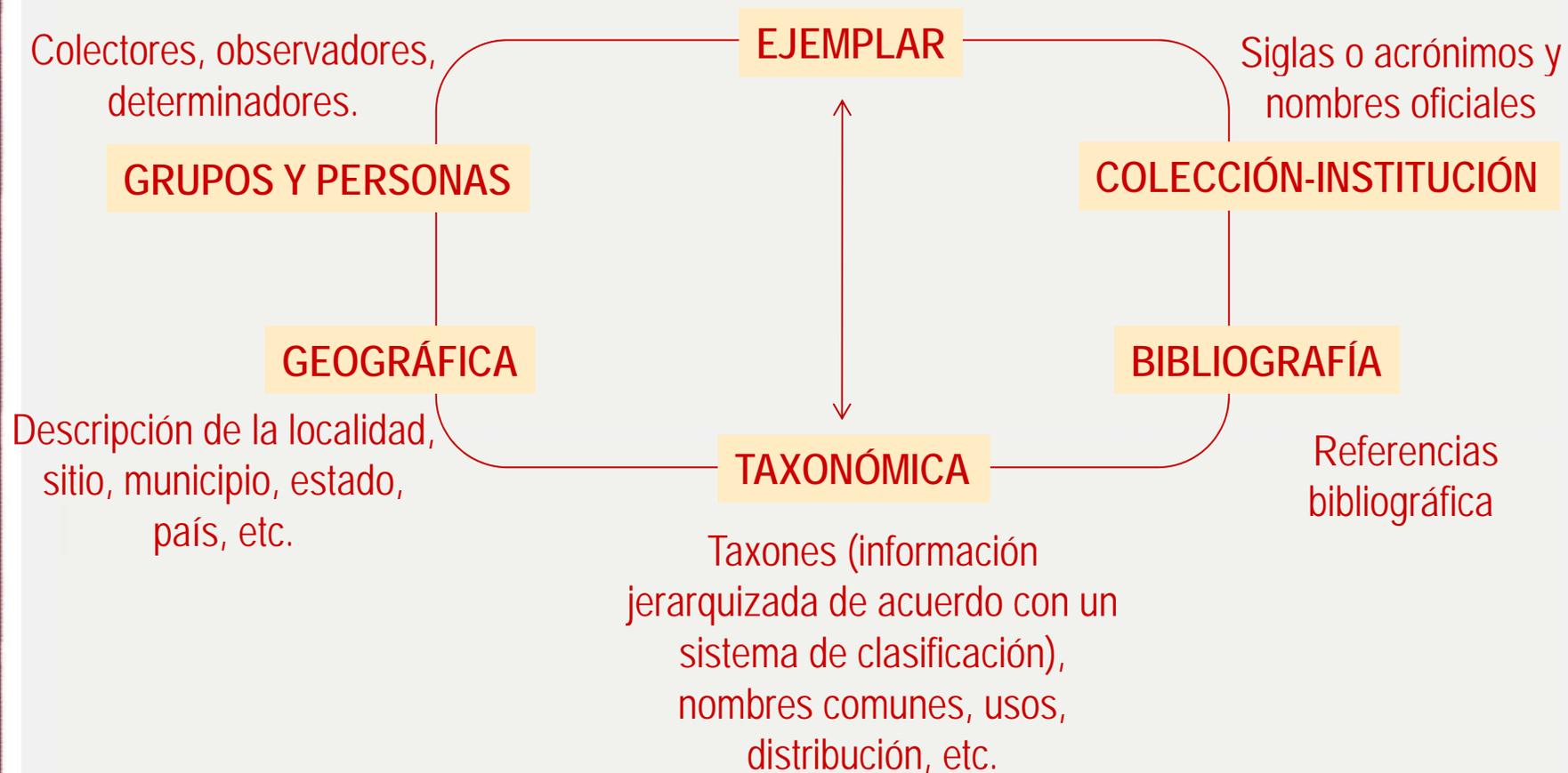


CONABIO

GOBIERNO FEDERAL

# Capas de información

Datos curatoriales o de observaciones.



## Objetivo. Evaluar la consistencia de la información biológica contenida en la base de datos

### Cómo:

- Mediante la revisión de la información contenida en las bases de datos (1... n veces).

### Con qué:

- Utilizando un método para evaluar la consistencia
  - ✓ Instructivo para la conformación de bases de datos
  - ✓ Protocolo de control de calidad.
  - ✓ Herramientas informáticas de análisis de bases de datos.
  - ✓ Catálogos de autoridad (CAT).
  - ✓ Bibliografía especializada.
  - ✓ Consulta a especialistas.
  - ✓ Convenio suscrito (proyecto y términos de referencia).



CONABIO

Gobierno  
Federal

# Algunos ejemplos de errores detectados en el control de calidad



Análisis automático

-> CG 2\_02 Error ortografico en region\_cat  
 -> CG 3\_01 Dato que no corresponde en region\_cat  
 -> CG 4\_01 Dato repetido en region\_cat  
 -> CG 5\_01 Dato sin convencion en region\_cat  
 -> CG 5\_01 Dato sin convencion en sitio  
 -> CG 5\_01 Dato sin regla de sintaxis en region\_cat  
 -> CG 5\_02 Nombre de tipos de region fuera de dominio  
 -> CG 5\_03 Valores incorrectos en estado o municipio\_INEGI  
 -> CG 5\_04 Ambiente sitio fuera del dominio  
 -> CG 5\_05 Obtencion coord fuera del dominio  
 -> CG 5\_06 Datos incorrectos en estado o municipio\_INEGI  
 -> CG 7\_04 Misma coordenada geografica con diferente municipio  
 -> CG 7\_06 Coordenadas incongruentes  
 -> CG 7\_07 Incongruencia entre altitud inferior y superior  
 -> CPG 1\_01 Dato faltante en grupos  
 -> CPG 1\_01 Dato faltante en personas  
 -> CPG 1\_03 Grupos sin integrantes  
 -> CPG 2\_01 Error tipografico 2,4 y 5 en grupos  
 -> CPG 2\_01 Error tipografico 2,4 y 5 en personas  
 -> CPG 4\_02 Grupos repetidos  
 -> CPG 5\_01 Grupos sin convenciones  
 -> CPG 5\_01 Personas sin convenciones  
 -> CTB 1\_01 Dato faltante campo AnotacionTaxon  
 -> CTB 1\_01 Dato faltante campo CitaNomenclatural  
 -> CTB 1\_01 Dato faltante campo Fuente  
 -> CTB 1\_01 Dato faltante campo SistClasCatDicc  
 -> CTB 1\_01 Dato faltante campo Taxon  
 -> CTB 1\_03 Dato faltante campo AutorAnio  
 -> CTB 1\_04 Taxones con estatus ND  
 -> CTB 1\_05 Taxones con estatus sinonimo sin asociar  
 -> CTB 1\_06 Especie e infraespecie con dato unico con ejemplares  
 -> CTB 1\_06 Especie e Infraespecie con dato unico sin ejemplares  
 -> CTB 1\_07 Especies e infraespecies nov con ejemplares

Seleccionar todos

Finalizado...

Control de calidad SNIB

Menú

Modelo SNIB

Base de datos

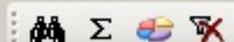
- Personas y grupos
- Ejemplar (curatoria)
- Taxonomica - biogeografica
- Geográfica
- Bibliográfica
- Colecciones e instalaciones
- Estadísticas
- Información externa
- Restricción de la información
- Sin capa

Sistema

able para

No

## Resultados del control de calidad



CE 1\_03 Ejems colect o report sin No de catalogo

IdEjemplar	CategoriaTaxonomica	Nombre	NumCatalogo	SiglasColeccion	ProcedenciaDatos	DescripcionProcedencia
973	especie	<i>Atrophecaecum astorquii</i>	Desconocido	CHCM	1	Colectado
1592	especie	<i>Neoechinorhynchus golvani</i>	NA	UJAT	1	Colectado
788	especie	<i>Procamallanus rebecae</i>	ND	UJAT	2	Reportado
771	especie	<i>Procamallanus rebecae</i>	ND	UJAT	2	Reportado
772	especie	<i>Procamallanus rebecae</i>	ND	CNHE	2	Reportado
773	especie	<i>Procamallanus rebecae</i>	ND	CNHE	2	Reportado
774	especie	<i>Procamallanus rebecae</i>	ND	CNHE	2	Reportado
775	especie	<i>Procamallanus rebecae</i>	ND	UJAT	2	Reportado
776	especie	<i>Procamallanus rebecae</i>	ND	UJAT	2	Reportado
777	especie	<i>Procamallanus rebecae</i>	ND	CNHE	2	Reportado
778	especie	<i>Procamallanus rebecae</i>	ND	UJAT	2	Reportado
779	especie	<i>Procamallanus rebecae</i>	ND	UJAT	2	Reportado
780	especie	<i>Procamallanus rebecae</i>	ND	UJAT	2	Reportado

### Campos Agrupados

TipoPreparacion	Totales
	2139
Preparación microscópica permanente	1
Alcohol 50%	1
Habitat estómago	7
kpal	1
Laminilla	1
nd	1
Nemátodos conser vados en frascos viales.	1
Nemátodos conservados en f rascos viales.	1
Nemátodos conservados en frascos viales.	360
planta acuática	5
Preparación microscópica permanente	1
Preparación microscópica permanente	1743

### Seleccionar errores

idError	Campo	Completo
680	TipoPreparacion	Ejemplar con dato de tipo de preparación que no corresponde
966	TipoPreparacion	Ejemplar con parte del dato de tipo de preparación que no corresponde

Seleccionar todos

## Resultados del control de calidad



CE 1\_10 Ejemplares sin determinar a especie

IdEjemplar	IdNombre	CategoriaTaxonomica	Nombre	Genero	Especie	Observaciones
23	25387	especie	Aphelandra sp.	Aphelandra	sp.	Ejemplar no determinado por ausencia de claves recientes.
37	25389	especie	Justicia sp.	Justicia	sp.	Ejemplar no determinado por ausencia de claves recientes.
47	25391	especie	Ruellia sp. 1	Ruellia	sp. 1	Ejemplar no determinado por ausencia de claves recientes.
48	25392	especie	Ruellia sp. 2	Ruellia	sp. 2	Ejemplar no determinado por ausencia de claves recientes.
50	538	familia	Acanthaceae	NP	NP	Ejemplar no determinado por ausencia de claves recientes.
51	25393	especie	Ruellia sp. 3	Ruellia	sp. 3	Ejemplar no determinado por ausencia de claves recientes.
52	25392	especie	Ruellia sp. 2	Ruellia	sp. 2	Ejemplar no determinado por ausencia de claves recientes.
54	25394	especie	Ruellia sp. 4	Ruellia	sp. 4	Ejemplar no determinado por ausencia de claves recientes.
55	538	familia	Acanthaceae	NP	NP	Ejemplar no determinado por ausencia de claves recientes.
56	25393	especie	Ruellia sp. 3	Ruellia	sp. 3	Ejemplar no determinado por ausencia de claves recientes.
57	25393	especie	Ruellia sp. 3	Ruellia	sp. 3	Ejemplar no determinado por ausencia de claves recientes.
89	25397	especie	Gomphrena sp. 1	Gomphrena	sp. 1	Ejemplar no determinado por ausencia de claves recientes.
90	25397	especie	Gomphrena sp. 1	Gomphrena	sp. 1	Ejemplar no determinado por ausencia de claves recientes.
100	25399	especie	Iresine sp. 1	Iresine	sp. 1	Ejemplar no determinado por ausencia de claves recientes.
101	25400	especie	Iresine sp. 2	Iresine	sp. 2	Ejemplar no determinado por ausencia de claves recientes.
103	25401	especie	Iresine sp. 3	Iresine	sp. 3	Ejemplar no determinado por ausencia de claves recientes.
104	25402	especie	Iresine sp. 4	Iresine	sp. 4	Ejemplar no determinado por ausencia de claves recientes.
105	25403	especie	Iresine sp. 5	Iresine	sp. 5	Ejemplar no determinado por ausencia de claves recientes.
116	98	familia	Amaranthaceae	NP	NP	Ejemplar no determinado por ausencia de claves recientes.
117	25406	especie	Pfaffia sp. 1	Pfaffia	sp. 1	Ejemplar no determinado por ausencia de claves recientes.



CONABIO

GOBIERNO  
FEDERAL

## Implementación de la herramienta de control de calidad plan de trabajo

- Herramienta de control de calidad y marcado de datos de posibles errores, inconsistencias y omisiones en el SNIB (avance 85%)
- Herramienta para comparar los nombres taxonómicos vs. catálogo de autoridad taxonómica (avance 85%)
- Compatible con Biótica y con el estándar DwC
- Probablemente desarrollaremos una herramienta para que el usuario haga el mapeo de datos y pueda utilizar la herramienta (ej. NoCatalogo, Catalogo, CatalogNumber, Número de catálogo)
- Disponible al público: base de datos, reporte de la calidad de los datos y herramientas informáticas



**GOBIERNO  
FEDERAL**

1992-2012  
**20 AÑOS  
CONABIO**



COMISIÓN NACIONAL PARA  
EL CONOCIMIENTO Y USO DE  
LA BIODIVERSIDAD



**¡Gracias!**

Dirección General de Bioinformática  
Subdirección de Inventarios Bióticos

**[sib@conabio.gob.mx](mailto:sib@conabio.gob.mx)**

