Jornada de la Red de Responsables Técnicos en Visibilidad Web UNAM

Buenas prácticas y consideraciones para repositorios digitales de datos

> Joaquín Giménez Héau Coordinador Técnico CCUD















- La UNAM cuenta con un vasto patrimonio de información primaria generada por el trabajo cotidiano de su personal académico.
- Este acervo es extraordinariamente extenso en su cobertura temporal, espacial y temática.
- Se encuentra almacenado en las diversas colecciones que custodian las entidades académicas de la UNAM.

















Coordinación de Colecciones **Universitarias Digitales CCUD**

Es una coordinación adscrita Secretaría de Desarrollo Institucional de la UNAM, con el fin de desarrollar una plataforma transdisciplinaria de datos, con herramientas para la integración, administración, consulta y difusión de los acervos digitales universitarios.





ACUERDO POR EL QUE SE CREA LA COORDINACIÓN DE COLECCIONES UNIVERSITARIAS DIGITALES

Autónoma de México, con fundamento en los artículos 1º las facultades y escuelas; y 9º de la Ley Orgánica; 34, fracciones IX y X del Estatuto

CONSIDERANDO

Que la Universidad Nacional Autónoma de México cuenta con un vasto patrimonio conformado por un extenso acervo de colecciones científicas, humanísticas y culturales custodiadas por las entidades que conforman los tres subsistemas, así como las facultades y escuelas.

Que las colecciones en su conjunto constituyen una herramienta invaluable para ampliar el conocimiento en temas globales y profundizar en la investigación de frontera en todas las áreas del saber.

Que, entre sus prioridades, la UNAM ha asumido de forma invariable la de hacer accesible a su comunidad, a la sociedad y al más amplio público que sea posible, los productos del conocimiento y los bienes que resultan de la investigación científica y humanística, así como de la expresión artística.

Que el empleo de las tecnologías de la comunicación y de la información es hoy en día indispensable para garantizar la preservación y divulgación de los bienes generados por la ciencia, las humanidades y las artes, así como para extraer e integrar de formas novedosas el conocimiento que almacenan

tancia integrar, organizar y sistematizar a través de una red tecnológica transdisciplinaria el acervo de colecciones que resquarda la UNAM, así como las que genera mediante el trabajo académico cotidiano.

responsabilidades la de diseñar, impulsar y coordinar proyectos académicos innovadores para la institución.

Que, en virtud del volumen y de la diversidad del acervo de colecciones que resquarda la UNAM, resulta conveniente la creación de una instancia universitaria responsable de las tareas que implica la detección, concentración, organización, sistematización, digitalización y divulgación de este natrimonio

En razón de lo anterior, he tenido a bien expedir el siguiente:

PRIMERO. Se crea la Coordinación de Colecciones Universita rias Digitales (CCUD), la cual guedará adscrita a la Secretaría Único. El presente Acuerdo entrará en vigor a la fecha de su de Desarrollo Institucional

Segundo. Los objetivos de la CCUD son los siguientes:

. Implementar una ciber-infraestructura que facilite la integración, interacción y aprovechamiento de las colecciones existentes en la UNAM, bajo resguardo de las entidades que forman los subsistemas de la investigación

Dr. José Narro Robles, Rector de la Universidad Nacional científica, de humanidades, de difusión cultural, así como

II. Recibir y sistematizar la información sobre las nuevas colecciones científicas, humanísticas y artísticas que de manera cotidiana genera el trabajo académico de la UNAM:

III. Definir estándares y protocolos para el manejo digital de las colecciones existentes en la UNAM, que garanticen su interoperabilidad;

IV. Proveer herramientas para la investigación transdisciplinaria mediante el acceso a las colecciones y su información

V. Difundir, a través de los medios adecuados a cada caso, las colecciones existentes en la UNAM.

TERCERO. La CCUD tiene las siguientes funciones:

I. Evaluar y perfeccionar de manera permanente el funcionamiento de las actividades de inventario y difusión de las colecciones existentes en la UNAM:

II. Vincularse con las entidades y dependencias universitarias que resguardan y generan colecciones;

III. Trabajar en colaboración con las instancias correspondientes para determinar, según la colección de que se trate, los formatos, medios y recursos tecnológicos adecuados para su

IV. Actualizar de manera continua el acervo digitalizado y ponerlo a disposición de la comunidad y del público interesado, a través del Programa "Toda la UNAM en Línea":

V. Rendir el informe anual, cuantitativo y cualitativo, de las Que la Secretaría de Desarrollo Institucional tiene entre sus actividades realizadas, así como formular su plan de trabajo;

> VI. Expedir, modificar y actualizar su reglamento interno y su normatividad operativa, y

> VII. Las demás que le encomiende el Secretario de Desarrollo Institucional.

> Cuarto. El titular de la CCUD será nombrado por el Secretario de Desarrollo Institucional previo acuerdo con el Rector.

> Quivro. La estructura y organigrama de la CCUD serán propuestos por su titular y definidos por el Secretario de Desarrollo. Institucional, previo acuerdo con el Rector.

publicación en Gaceta UNAM.

"POR MI RAZA HABLARÁ EL ESPÍRITU" Ciudad Universitaria, D.F., a 14 de enero de 2013

DR JOSÉ NARRO ROBLES

20 /aceta 14 de enero de 2013

Acuerdo por el que se crea la Coordinación de Colecciones Universitarias Digitales -Gaceta UNAM, 14 de enero de 2013.

La CCUD se enfoca en las versiones digitales de las colecciones universitarias, definidas como el conjunto de datos primarios de la misma clase, reunidos y clasificados por su valor académico, artístico o patrimonial.



Datos primarios: son datos recabados directamente de los hechos que se investigan.

Datos secundarios: son datos derivados de la interpretación o integración de los datos primarios.



Problemas comunes en las colecciones digitales

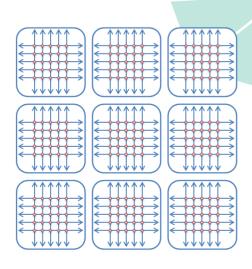
Visibilidad

Contexto

Integración







Los datos sólo pueden ser consultados a través de portales específicos. Forman parte de la web invisible y no pueden ser recuperados por motores de búsqueda como Google.

En general, las colecciones no cuentan con información sobre su estructura, antecedentes, características, valor, forma de uso y administración (p.e. autor, fecha de la última actualización, derechos).

Las colecciones no comparten plataforma ni estructura de datos, es decir, no están estandarizadas.



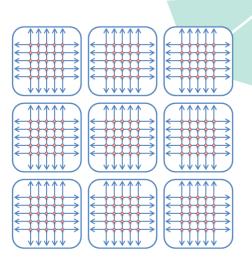
Visibilidad

Toda la

Contexto



Integración



El contenido de las bases de datos debe replicarse en formatos compatibles con los motores de búsqueda y con internet en general, a través de servicios web. Esto permite aprovechar los sistemas de búsqueda e indexación que existen en la red.

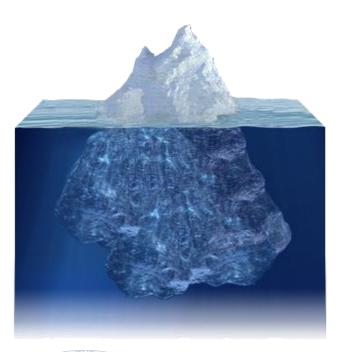


Buenas prácticas

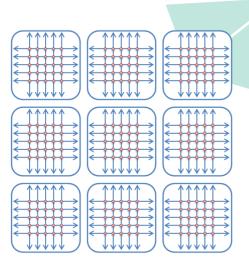
Visibilidad

Contexto

Integración











Cada registro que se consulte desde internet debe incluir la información que describe a la colección así como su estructura y relación con otras bases de datos.



Buenas prácticas

Visibilidad



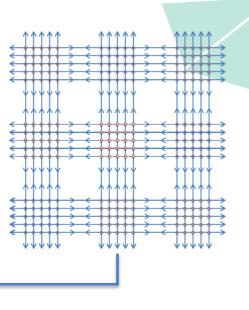




Contexto







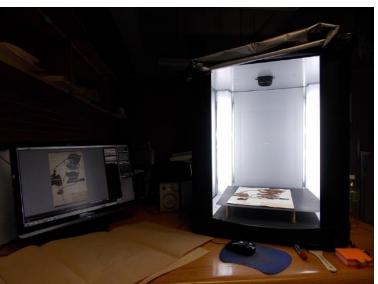
Debe haber un proceso permanente de investigación y documentación para definir estructuras estandarizadas de datos, además de integrar tecnologías de indexación (Apache-Solr-Lucene), etiquetado de datos y recursos, y redes de servicios web.



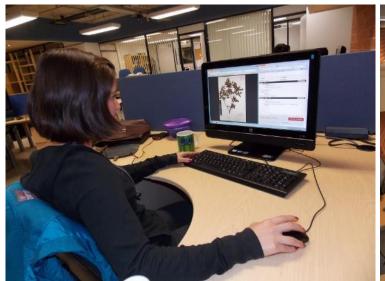






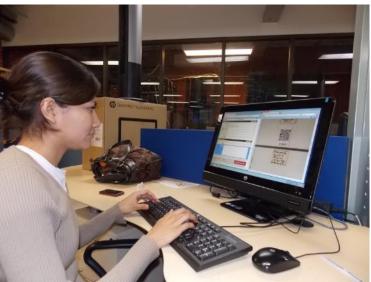






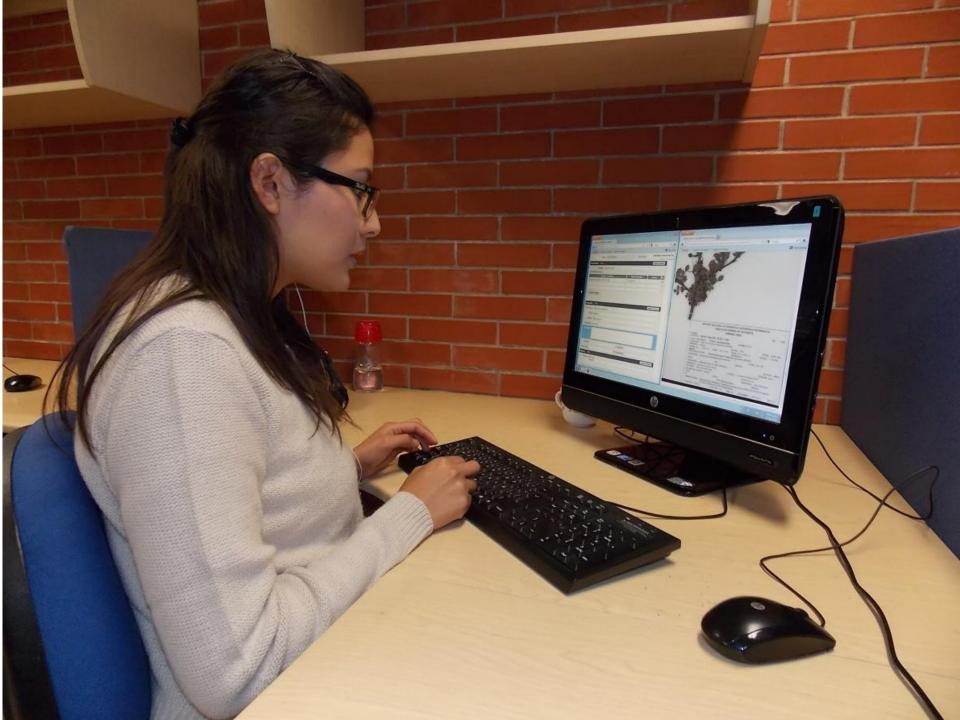
















670- Cordia sp. No hay fulos Borragináceas Prof. C. Conrath

BORAGINACEAE DE MÉXICO

BORAGINACEAE DE MÉXICO Varronia globosa Jacq.
Ma.Goreti Campos-Rios (CICY), CONABIO GE025, 2010

Cordia globosa (Jacq.) H.B.K. Determined by James S. Miller 1986
Missouti Botanical Garden

N. v. Casthering N. Procedencia de la companya de la

Colocto Willey

670- Cordia sp. No hay frutos Borraginaceas Prof. C. Conkathi Min. 670.

Procedencia. Colecto Minu

BORAGINACEAE DE MÉXICO

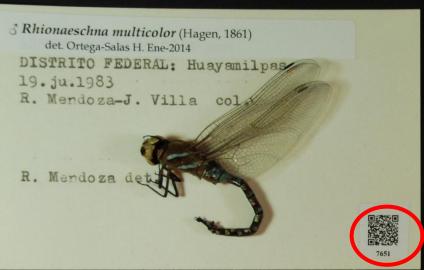
BORAGINACEAE DE MÉXICO

Varronia globosa Jacq.

Ma. Goreti Campos-Ríos (CICY), CONABIO GEO 5, 2010

Cordia globosa (Jacq.) H.B.

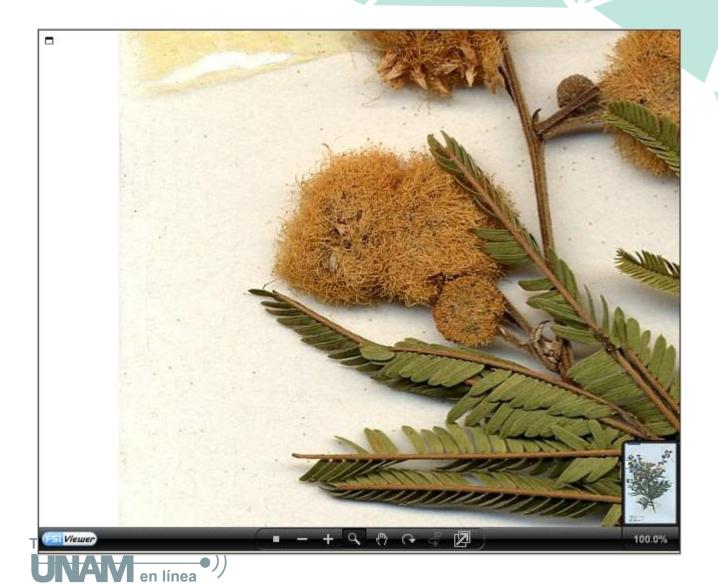
Determined by James S. Miller 1986 Missouri Botanical Garden























Abies vejarii

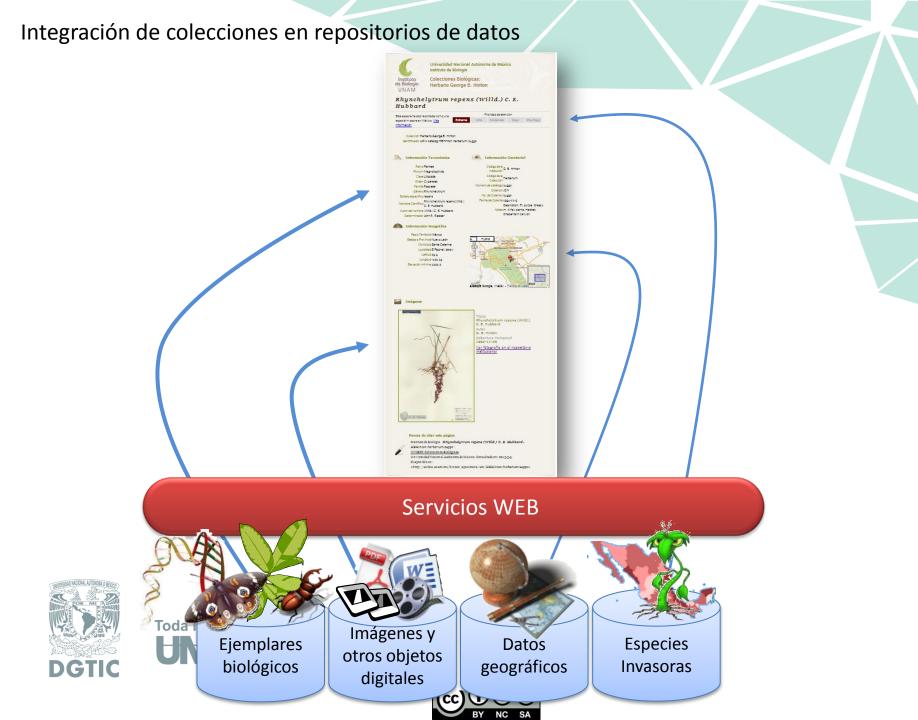


Folio: 701932	Verificar	Herbario Nacional de Méxic Plantas Vasculares
Taxonomía <f5></f5>		repetir antenor
Familia: Piperaceae		
Determinaciones:		
Nombre Científico	Determinador	Fecha Últ.
Piper pseudo-fuligineum C. de Candolle	Ursula Bachem C.	O CONTROL OF THE PARTY OF THE P
		SERVING MICHAEL
		MEXICO MEXICO
Añadir renglón		
Geografia <f7></f7>		resets anterior
Ubicacións		
País: México		
Estado: Chiapas		
Municipio: Villa Corzo		
Localidad: Predio Particular Europa.		H, 5
		-
Latitud: L	.ongitud:	
Elevación: 1350 msnm Pro	fundidad:	
Colecta <f6></f6>		repath unterfor
Fecha y hora de		
colecta: 07/06/1989		
Colector: Ursula Bachem C. y Ric	cardo Rojas C.	
Número de colecta: 853		791932
Método de colecta:		
Duplicados: 1		
Tipo de preparación: Herborizado		HERRADO DE LA DIVIDION DE CIERCIAS POREITALAS. (NEVERSIDAD APPONDA CHAPTONIA (CHAPTON)
Biológicos y otros <f8></f8>		PLANTAG DE DEPRESOR CENTRAL Y HIFBIA MACHE DE CIESAFAG
Forma de vida:		Piper provide railstanum C. CC. PYPERFOR Confusesite
Hábitat: Bosque mesófilo de me	ontaña con Inga, Eupat	
Información Fanerófita /// 4 m de		
adicional: inconspicuas /// 2 cm escasa	de diámetro /// Freci	Francista. A. CO is the Alberts, Elementation of the Commissional Comm
		Contractor, marine.
		The state of the s
		teri, tenter finites (- y Rinardo bejon C. 983) teri, Dreifa Beden C.

















Esta especie ha sido reportada como una especie invasora en México. Más información Extrema Alta Moderada Baja Muy Baja













<u>:10</u>

Rhynchelytrum repens

Prioridad de atención:

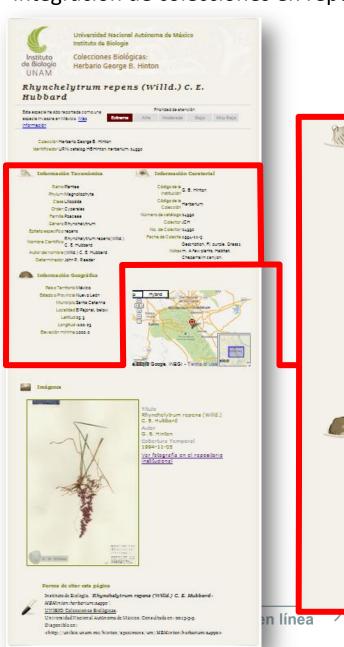
Reino	<u>Plantae</u>
División	Magnoliophyta
Clase	<u>Liliopsida</u>
Subclase	Commelinidae
Orden	Cyperales
Familia	<u>Poaceae</u>
Género	Rhynchelytrum
Especie	repens
Categoría	sp.

Prioridad de atención:	
	Extrema

Índice de invasividad por Estado









Información Geográfica

País o Territorio México

Estado o Provincia Nuevo León

Municipio Santa Catarina

Localidad El Pajonal, below

Latitud 25.3

Longitud -100.23

Elevación mínima 1000.0

CC (1 S) O BY NC SA Información Curatorial

Código de la Institución G. B. Hinton

Código de la Colección Herbarium

Número de catálogo 24990

Colector JCH

No. de Colector 24990

Fecha de Colecta 1994-11-5

Description: Fl. purple. Grass 1

Notas m. A few plants. Habitat:
Chaparral in canyon.

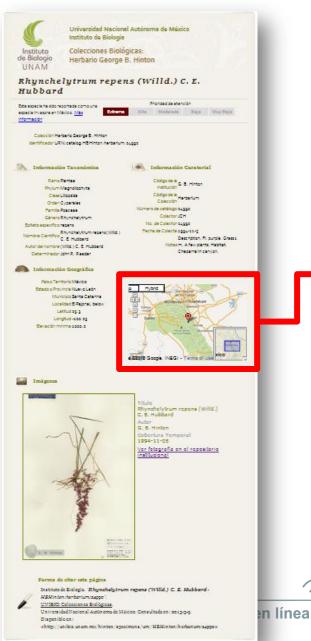
POACEAE
Rhynchelytrum repens (Willd.) C. B.
Hubbard

Determined by J. R. Reeder

HKRBARIUM OF G. B. HINTON NO. 24990
Collected by Hinton et al. 1994/11/05
El Pajonal, below. 1000 m.
Santa Catarina, N. L., México.

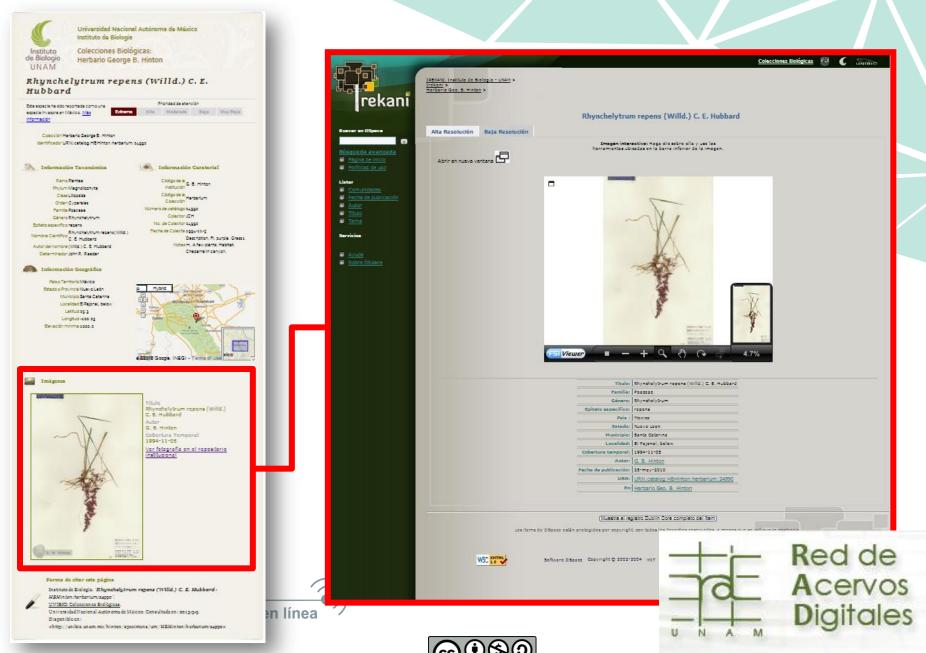
Fl. purple. Grass 1 m. A few plants.

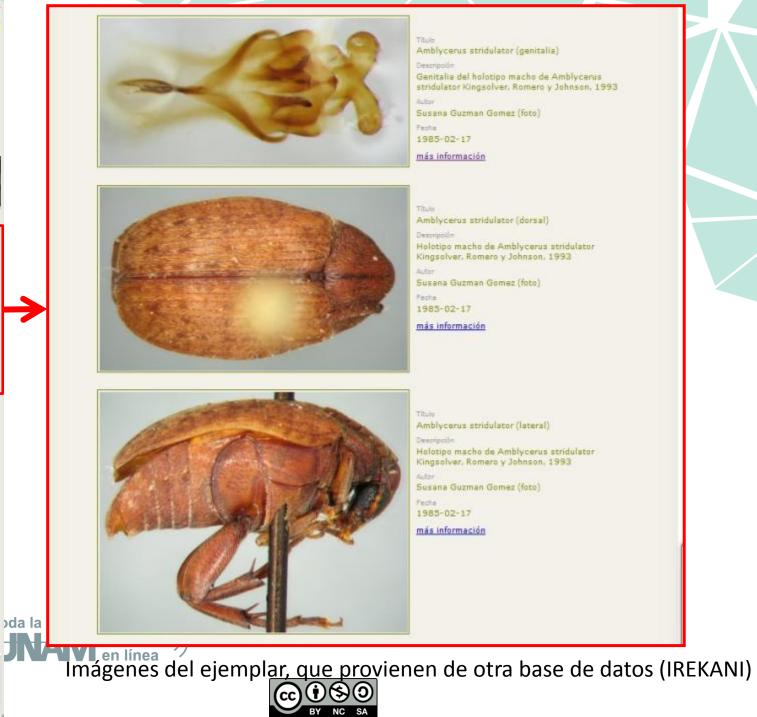
Chaparral in canyon.

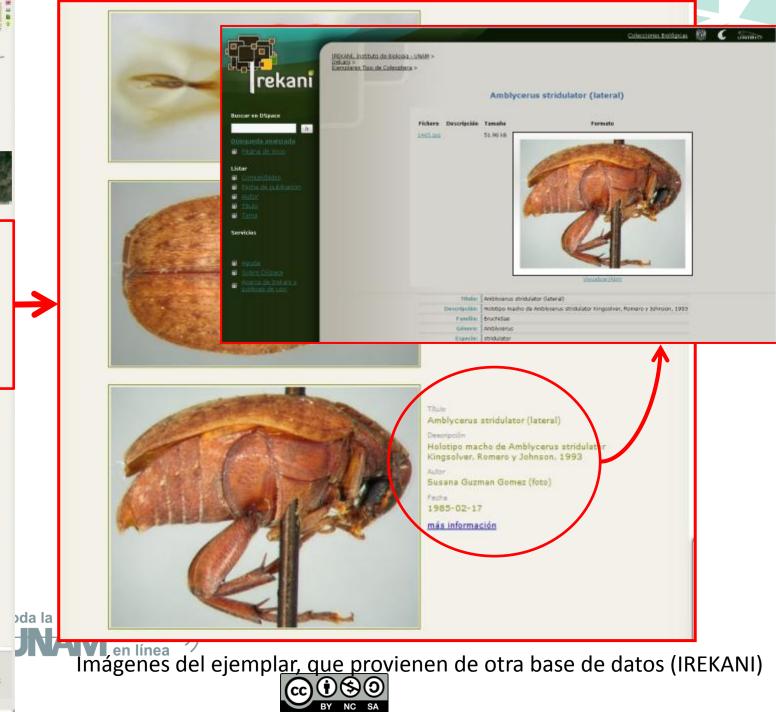






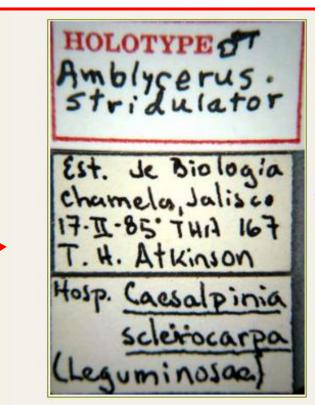








Se pueden almacenar imágenes de las etiquetas, con el fin de preservar los datos originales



Amblycerus stridulator (etiquetas)

Descripción

Etiquetas del holotipo macho de Amblycerus stridulator Kingsolver, Romero y Johnson, 1993

Autor

Susana Guzman Gomez (foto)

Fecha

1985-02-17

más información











n linea

Forma de citar esta página

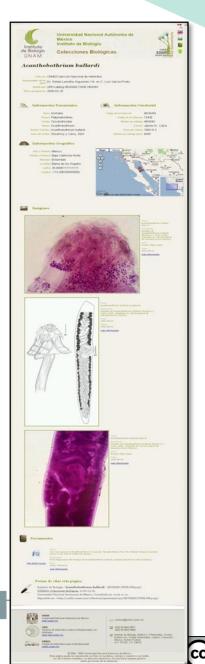
Instituto de Biología. "Rhynchelytrum repens (Willd.) C. E. Hubbard - HBHinton:herbarium:24990". UNIBIO: Colecciones Biológicas.

Universidad Nacional Autónoma de México. Consultada en: 2013-3-5.

Disponible en: http://unibio.unam.mx/hinton/specimens/urn/HBHinton:herbarium:24990



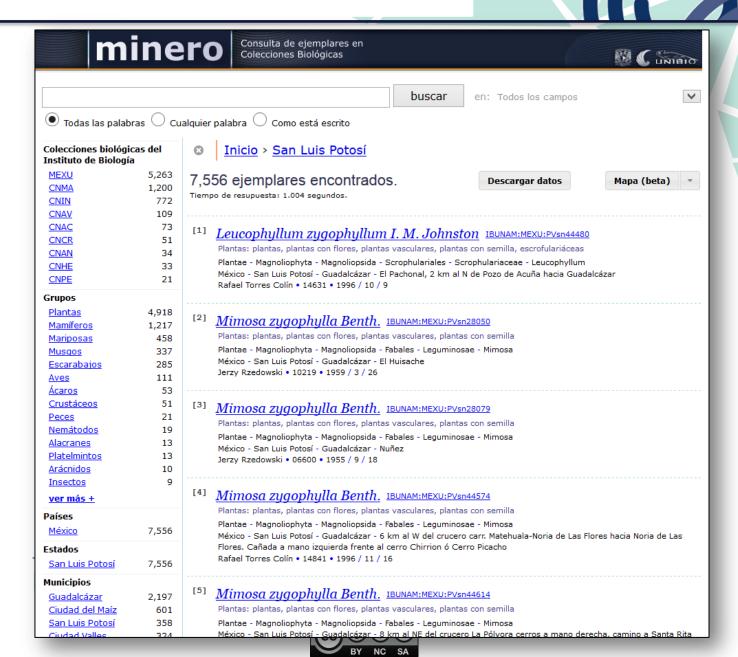








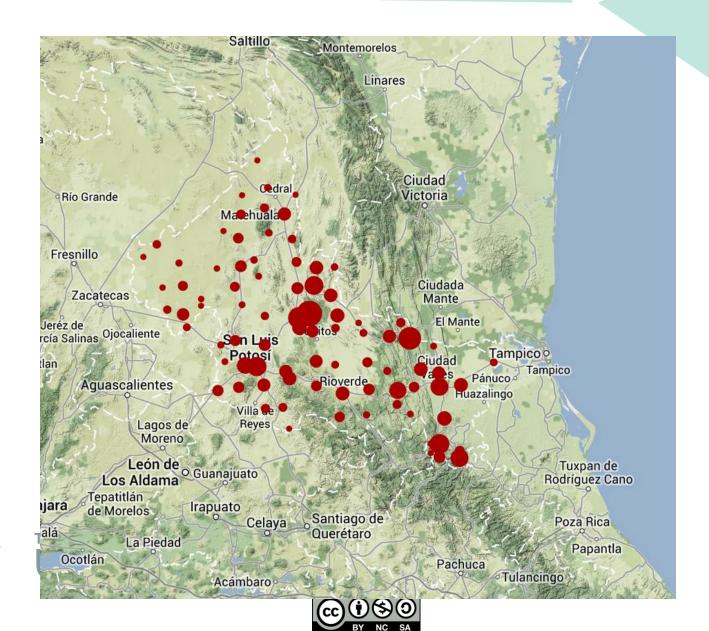
Consulta de datos



CCUD



Visualización de datos





Tipos de información en las colecciones digitales universitarias

Registros en colecciones

Conjunto de datos correspondientes a un registro de una base de datos que tienen valor por sí mismos. (p. ej. <u>Ejemplar de herbario</u>).

- Pueden publicarse como fichas.
 Pueden seguir estándares.
 - Pueden tener metadatos.

Capas geoespaciales (Mapas)

Conjunto de datos con referencia geoespacial que tienen valor en un contexto geográfico (p. ej. Mapa de Áreas Naturales Protegidas).

- No pueden publicarse como fichas.
- Pueden seguir estándares.
- Pueden tener metadatos.

Bases de datos



Conjunto de datos que se encuentran contenidos dentro de un manejador de bases de datos o tabla y que sólo tienen valor dentro de ese contexto (p. ej. Mediciones de temperatura en un lago).

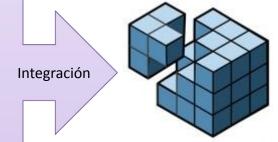
- No pueden publicarse como fichas.
- Pueden seguir estándares.
- Pueden tener metadatos.

Objetos digitales

Conjunto de datos indivisible que en su totalidad y estructura conforman una representación. Tiene valor por sí mismo y sólo pueden manejarse a nivel de metadatos (p. ej. Una imagen).

- Pueden publicarse como fichas (metadatos).
- Pueden seguir estándares (metadatos).
- Pueden tener metadatos.

Interoperabilidad



Construir sistemas complejos y grandes





DATOS ABIERTOS

Los datos abiertos son datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona

- **1. Disponibilidad y acceso:** la información debe estar disponible descargándola de internet, en una forma conveniente y modificable.
- 2. Reutilización y redistribución: los datos deben ser provistos bajo términos que permitan reutilizarlos y redistribuirlos, e incluso integrarlos con otros conjuntos de datos.

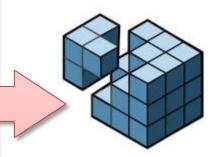
CC BY-SA

BY - Atribución, requiere la referencia al autor original. SA - Compartir Igual, permite obras derivadas bajo la misma licencia o similar (posterior u otra versión por estar en distinta jurisdicción).



3. Participación universal: todos deben poder utilizar, reutilizar y redistribuir la información. No debe haber discriminación alguna en términos de esfuerzo, personas o grupos. Sin restricciones comerciales o de uso bajo ciertos términos.

Interoperabilidad



Construir sistemas complejos y grandes



Formas de presentar los datos

Disponible en la Internet (en cualquier formato. Por ejemplo: PDF), siempre que sea con licencia abierta, para que sea considerado un Dato Abierto.	PDF PDF Adobe	
Disponible en la Internet de manera estructurada (en un archivo Excel con extensión XLS)	Excel	
Disponible en la Internet de manera estructurada y en formato no propietario (CVS en vez del Excel).	Texto separado por comas	
Siguiendo todas las reglas anteriores, pero dentro de los estándares establecidos por el W3C (RDF e SPARQL): usar URL para identificar cosas y propiedades, de manera que las personas las puedan direccionar para sus publicaciones.	Triplete http://www	
Todas las reglas ya mencionadas, y además: vincular sus datos a los de otras personas, de manera a proveer un contexto.	LINKED DATA	***





Linked data es un método de publicación de datos estructurados para que puedan ser interconectados con otros datos de manera automática.



Integración de colecciones (linked data)



Esta especie ha sido reportada como una especie invasora en México. Más información Extrema Alta Moderada Baja Muy Baja









Monitoreo del estado de las invasiones biológicas de plantas en México



Rhynchelytrum repens

Drincidad de atención:

Reino Plantae
División Magnoliophyta
Clase Liliopsida
Subclase Commelinidae
Orden Cyperales
Familia Poaceae
Género Rhynchelytrum
Especie repens
Categoria 5p.

Filoridad de aterición.	
	Extrema

Índice de invasividad por Estado



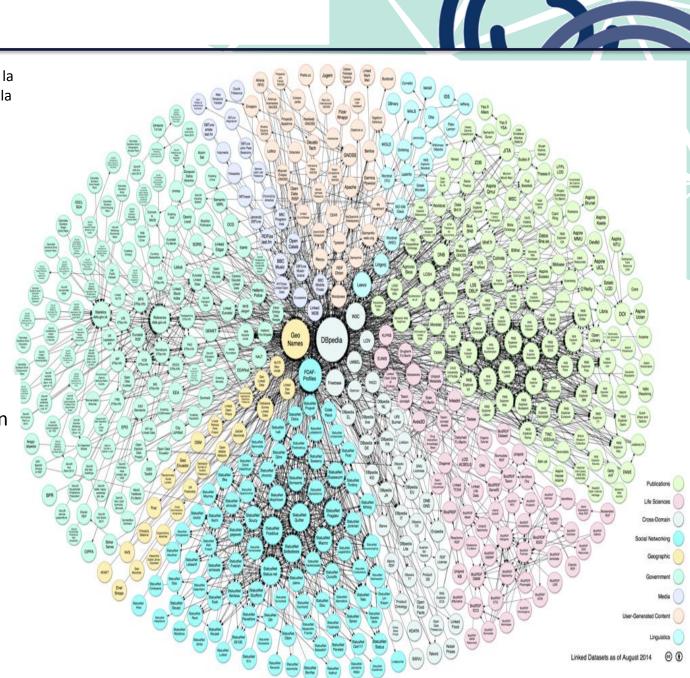


Proyecto desarrollado por el grupo de la W3C encargado de divulgar y explicar la Web semántica (Semantic Web Education and Outreach)

Busca ampliar la web con una base de datos común mediante la publicación en la Web de bases de datos en RDF y mediante el establecimiento de enlaces RDF entre datos de diferentes fuentes.

RDF es el Marco de

Descripción de Recursos es un
método general para la
descripción conceptual o
modelado de la información
que se implementa en los
recursos web, utilizando una
variedad de notaciones de
sintaxis y formatos de
serialización de datos.





¡Gracias!

joaquin@ccud.unam.mx

www.ccud.unam.mx

contacto@ccud.unam.mx

Teléfono: 56228971

