

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### Índice

1. Introducción
2. Antecedentes
3. Unificación de criterios para la captura de los datos
4. Interfaz de captura UNIBIO
  - 4.1. Elementos de la Interfaz de captura
  - 4.2. Interfaz de Reportes de problemas en ejemplares
  - 4.3. Notas sobre la Interfaz de captura
5. Recepción de ejemplares e imágenes
  - 5.1. Problemas en la recepción de ejemplares e imágenes
6. Ingreso de registros
7. Reglas de captura
  - 7.1. Reglas generales
    - 7.1.1. Captura total
      - 7.1.1.1. Diversidad de etiquetas
        - 7.1.1.1.1. Etiqueta principal
        - 7.1.1.1.2. Etiquetas secundarias
          - 7.1.1.1.2.1. Adicional al evento
          - 7.1.1.1.2.2. Nota accesoria
          - 7.1.1.1.2.3. Texto asociado
          - 7.1.1.1.2.4. Evento de determinación
    - 7.1.2. Asignación y atomización de la información
      - 7.1.2.1. Asignación
        - 7.1.2.1.1. Asignación de información asociada
      - 7.1.2.2. Atomización
      - 7.1.2.3. Doble dato
    - 7.1.3. Captura verbatim
      - 7.1.3.1. Sin formato de control
        - 7.1.3.1.1. Error de origen
        - 7.1.3.1.2. Probable error
        - 7.1.3.1.3. Variante
        - 7.1.3.1.4. Mayúsculas



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

- 7.1.3.1.5. Signos de puntuación
- 7.1.3.1.6. Denominador de campo
- 7.1.3.1.7. En dos idiomas
- 7.1.3.2. Con formato de control
  - 7.1.3.2.1. Escrita a mano
  - 7.1.3.2.2. Ilegible
  - 7.1.3.2.3. Extranjera
  - 7.1.3.2.4. Confusa
  - 7.1.3.2.5. Opciones múltiples, de llenado o marcado
  - 7.1.3.2.6. Múltiples
  - 7.1.3.2.7. Ortotipografía y correcciones
  - 7.1.3.2.8. Enmendadura
  - 7.1.3.2.9. Dañada

### 7.2. Reglas específicas (campos)

- 7.2.1. Familia
  - 7.2.1.1. Varias familias
- 7.2.2. Nombre científico
  - 7.2.2.1. Nombre científico más notas taxonómicas
  - 7.2.2.2. Diferentes eventos de determinación taxonómica
  - 7.2.2.3. Confirmación de nombre científico
  - 7.2.2.4. Números asociados a taxa
- 7.2.3. Determinadores y colectores
  - 7.2.3.1. Firma
- 7.2.4. Fechas de determinación y colecta
- 7.2.5. Número de colecta
- 7.2.6. País, estado, municipio y localidad
- 7.2.7. Coordenadas
- 7.2.8. Elevación
- 7.2.9. Forma de vida
- 7.2.10. Hábitat
- 7.2.11. Características del ejemplar
- 7.2.12. Nombre común



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

Sugerencia de cita: DGRU.2020. Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2 (Versión 1). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.13. Información adicional

- 7.2.13.1. Notas de campo
- 7.2.13.2. Factores ambientales
- 7.2.13.3. Valores promedio anual o total
- 7.2.13.4. Clima
- 7.2.13.5. Condición reproductiva
- 7.2.13.6. Taxa asociados
- 7.2.13.7. Abundancia
- 7.2.13.8. Usos
- 7.2.13.9. Medios de establecimiento
- 7.2.13.10. Hábito y comportamiento
- 7.2.13.11. Encabezados y pies de etiquetas principales y secundarias
- 7.2.13.12. Sellos
- 7.2.13.13. Etiquetas con “Información asociada”
- 7.2.13.14. Etiquetas con “Texto asociado”
- 7.2.13.15. Etiquetas de “Nota accesoria”
- 7.2.13.16. Anotaciones de tipo curatorial
- 7.2.13.17. Números sin asociación

### 8. Problemas asociados a imágenes

- 8.1. Etiqueta doblada
- 8.2. Objeto superpuesto
- 8.3. Imagen incompleta

### 9. Problemas de captura

- 9.1. Abreviar
- 9.2. Adición
- 9.3. Asignación
- 9.4. Asignación total
- 9.5. Completar
- 9.6. Digitación
- 9.7. Incompleta
- 9.8. Procedimiento
- 9.9. Omisión





## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

---

- 10. Bibliografía
- 11. Anexos



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 1. Introducción

Esta guía tiene por objetivo servir de modelo para la captura eficiente de la información de las colecciones del Instituto de Biología, de la Universidad Nacional Autónoma de México (IBUNAM).

Esta guía es resultado del análisis y depuración de los datos capturados hasta el momento de la publicación de este documento, realizado por los analistas de la Dirección General de Repositorios (DGRU). Así mismo, fue complementada con las respuestas, comentarios y opiniones procedentes del cuestionario aplicado a los capturistas de la Unidad de Informática para la Biodiversidad (UNIBIO): “Análisis exploratorio del proceso de captura de los datos de las colecciones del Instituto de Biología, UNAM”.

Gracias al trabajo realizado se han podido detectar algunos errores e inconsistencias recurrentes. Estos errores pueden ser solucionados y evitados con lo que se mejora el resultado de digitalización de las colecciones que se realiza.

En esta guía se hace una descripción de los problemas de captura de datos más comunes y se describe su fuente de origen. También, se realizan sugerencias para evitarlos y con ello establecer la unificación de criterios de entrada de datos a la Interfaz de captura de la UNIBIO.

Este ejercicio pretende ser acumulativo, mejorable y sujeto a sugerencias, por lo que se exhorta a emitir comentarios sobre el mismo, así como del proceso de captura de la información de las colecciones del Instituto de Biología, UNAM.

### 2. Antecedentes

De acuerdo con algunos análisis sobre la calidad de la información en bases de datos de diversos ámbitos, entre el 60% y 90% de los datos presentan problemas de calidad (Dasu, Vesonder & Wright, 2003). La firma Gartner (2007) indica que en el ámbito empresarial un 25% de cualquier compañía presenta errores.

Para solucionar estos problemas es necesario implementar un proceso de control de calidad que permita que los datos sean óptimos, accesibles, precisos, oportunos, completos, coherentes, pertinentes, integrales, detallados, comprensibles, congruentes con otras fuentes, asimismo deben ser fáciles de leer y de interpretar (Fürber & Hepp, 2011; Redman, 2001; Singh & Singh, 2010).

En el control de calidad de datos existen varias técnicas, que van desde la detección hasta la solución de errores. En primer lugar, la prevención de errores es uno de los procesos que está estrechamente relacionado tanto con la colección de la que proceden como con su entrada en una base de datos (Chapman, 2005). De acuerdo con Maletic y Marcus (2000), aunque el esfuerzo de prevención de errores sea considerable, el hecho es que en grandes conjuntos de datos los errores seguirán existiendo, por lo que procesos como la validación y corrección de los datos son necesarios. La prevención de errores se considera superior a la detección y corrección de errores, ya que la detección es a menudo costosa y no puede garantizar un 100% de éxito (Dalcin, 2004).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

Sugerencia de cita: DGRU.2020. Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2 (Versión 1). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Por otro lado, la detección de errores tiene un papel especialmente importante cuando se trata de colecciones heredadas (Chapman & Busby, 1994; Dalcin, 2004; English, 1999). Como es el caso de gran parte de los datos primarios de las colecciones biológicas.

Otro de los elementos de control de calidad fundamentales en los procesos de digitalización de datos es la preservación de los datos originales. Estos datos comprenden, principalmente, los registrados por el colector y los añadidos más tarde por los curadores (Chapman, 2005), aunque también pueden considerarse los que son agregados por el personal involucrado en el mantenimiento y revisión de los ejemplares.

La preservación de la información íntegra de los ejemplares permitirá realizar el diagnóstico del estado de los datos de las colecciones del Instituto de Biología, UNAM, así como registrar la historia de manejo de los ejemplares (p. ej., la actualización de los datos taxonómicos). Al mantener los datos originales y llevar a cabo su depuración, si se cometen errores, la información original se puede recuperar (Chapman, 2005).

Por lo que para mantener los datos verdaderos e íntegros contenidos de las etiquetas se deben conservar y capturar de manera verbatim, es decir, que deben ser capturados palabra por palabra en la misma forma y orden en que aparecen en las etiquetas. Esto incluye cualquier falta ortográfica, abreviaturas, inconsistencias, etc. (Tegelberg, Haapala, Mononen, Pajari, & Saarenmaa, 2012).

Dado que la limpieza, estandarización y actualización de los datos son acciones posteriores a la captura, éstas deben ser documentadas detalladamente y para ello es necesario mantener un registro adecuado de dichos procesos (Chapman, 2005). Esta documentación también ayuda a los curadores y custodios de datos para realizar un seguimiento de los datos y su calidad y no desperdiciar recursos en una segunda verificación de supuestos errores (Chapman, 2005).

Por tal razón, la captura de datos debe ser de manera verbatim, ya que esta es una práctica que permite prevenir errores y posibilita la diagnosis de la calidad estructural de los datos originales. También es necesario considerar que cuando la información ha sido modificada, esta puede ser difícil o imposible de recuperar, por lo que puede generar problemas adicionales (Chapman, 2005), por lo tanto el seguimiento de cambios en la información dentro de la base de datos es tan imperante como mantener los datos originales.

### 3. Unificación de criterios para la captura de los datos

Gracias al análisis realizado sobre los datos de las colecciones del IBUNAM capturados, se ha podido registrar y categorizar distintos casos y tipos de errores en la información. Con esto se ha desarrollado un protocolo de unificación de criterios de la captura, con la intención de reducir, e incluso eliminar la incidencia de errores e inconsistencias similares en el futuro.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Es necesario precisar que los derechos y obligaciones que se tienen sobre el manejo de los datos en la cadena de digitalización de los datos de las colecciones del IBUNAM son diferentes de acuerdo a la etapa que se trate. Por lo que no todos los cambios a realizar se encuentran dentro de los alcances de la DGRU y como algunos únicamente les corresponden a los curadores, se hace presente, nuevamente, la importancia de la captura verbatim, por lo que se debe evitar corregir o modificar los datos de los ejemplares al capturarlos.

A continuación, se realiza una revisión de las distintas fases que conforman el proceso de captura, así como una descripción de las condiciones de las fuentes de los datos (ejemplares y fotografías), instrucciones de captura y los errores comunes que se han identificado.

### 4. Interfaz de captura UNIBIO

La Interfaz de captura de la UNIBIO tiene características tales que la hacen fácil de usar, ya que evita la complejidad que otras plataformas presentan para la entrada de datos biológicos. Por ejemplo, no se requiere del uso obligatorio de catálogos al momento de la captura, lo que propicia que la captura se lleve a cabo en lapsos de tiempo cortos (Gernandt, Sánchez, Samper, Giménez, & Salazar, 2014).

Además esto permite monitorear, en tiempo real, las estadísticas de los datos capturados por usuario, al mismo tiempo se pueden generar informes sobre los errores o incongruencias presentes en los ejemplares. De tal modo que, la normalización y depuración de la información se lleva a cabo en etapas posteriores a la captura. Con lo cual los datos obtienen un estado óptimo para su publicación.

#### 4.1 Elementos de la Interfaz de captura

La Interfaz de captura se puede visualizar por medio de cualquier explorador de internet desde los equipos de cómputo en las instalaciones de UNIBIO. Para ingresar a ella, el responsable de la UNIBIO debe registrar a cada capturista y asignar un nombre de usuario y contraseña.

En la primera ventana de la Interfaz son solicitadas estas credenciales y la elección de la colección que se va a capturar. La segunda ventana contiene, en primer lugar, un campo para el folio o número de catálogo según el tipo de colección y un botón de verificación. Así como el nombre del usuario y la colección a la que se ingresa (Figura 1).

También se incluyen los campos principales, según Darwin Core 2, agrupados en cuatro secciones. La primera sección comprende los datos taxonómicos (familia, nombre científico, determinadores y fechas de determinación de los distintos eventos). La segunda corresponde a datos geográficos (país, estado, municipio, localidad, coordenadas, elevación y profundidad). La tercera sección comprende los datos relacionados con el evento de colecta (colector, fecha y número de colecta, entre otros) y



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

finalmente una cuarta sección denominada Biológicos y otros, en la cual se incluyen datos biológicos sobre el ejemplar colectado (forma de vida, hábitat, características del ejemplar) y otros datos más generales como nombre común e información adicional. Aparece también un campo para capturar el URL de la imagen, en caso de que ésta haya sido la fuente de los datos del ejemplar.

Adicionalmente, se incluye otra serie de elementos para el guardado (Guardar datos), adición de nuevos registros (Nuevo Registro) y control de la captura (contador de registros capturados e indicador de cambios); así como un botón de acceso a la Interfaz de reportes.

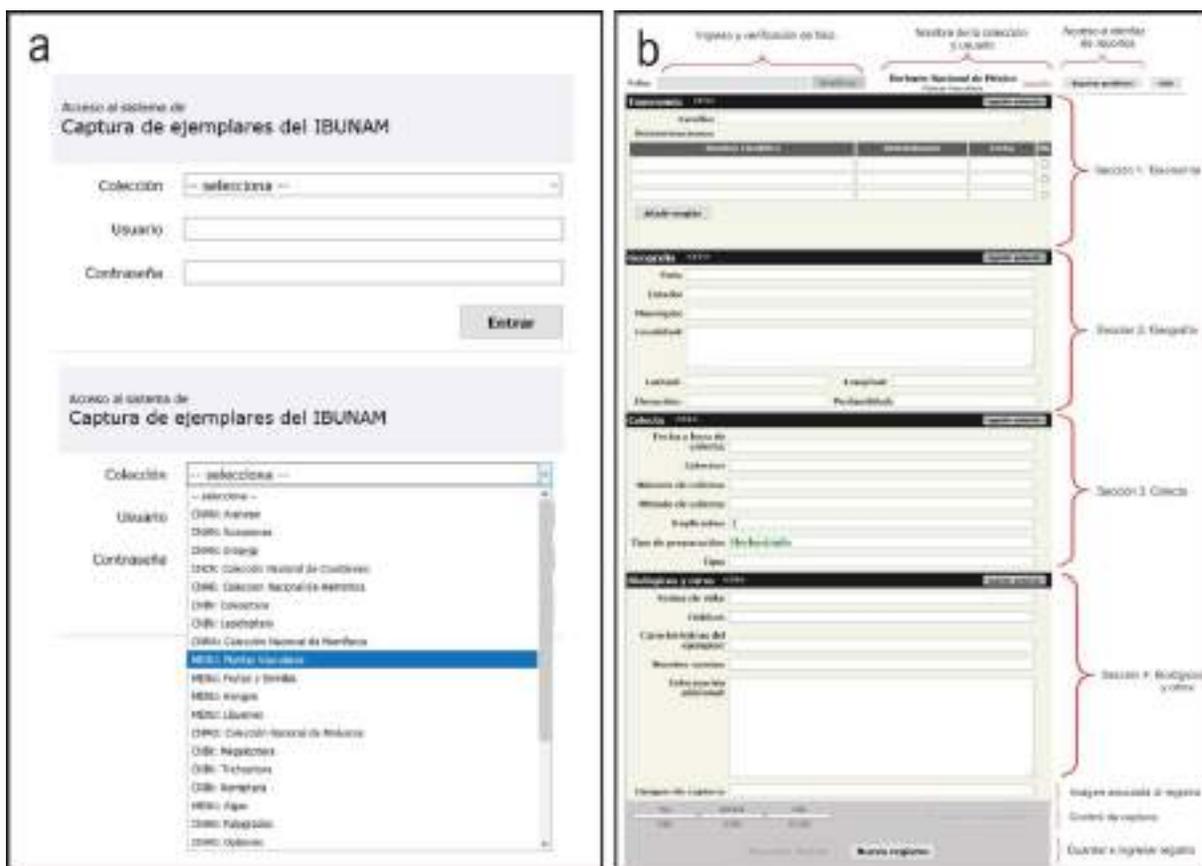


Figura 1. Interfaz de captura de la Unidad de Informática para la Biodiversidad (UNIBIO). a) Vista principal de ingreso a la Interfaz. Se muestran algunas de las colecciones registradas en la Interfaz; b) Secciones y elementos de la Interfaz de captura.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 4.2 Interfaz de Reportes de problemas en ejemplares

Dentro de la Interfaz de captura, se encuentra un botón que despliega la Interfaz de Reportes de problemas en ejemplares. Esta consta de una ventana emergente (Figura 2) en la cual se puede reportar cualquier problema relacionado con el estado de los ejemplares o sus imágenes. De igual forma se pueden reportar los problemas asociados a la información contenida en las etiquetas de los ejemplares (físicos o fotografías). Los tipos de problemas que deben ser reportados por medio de esta Interfaz serán descritos a detalle más adelante.



Figura 2. Interfaz de Reportes de problemas en ejemplares. Empleada en la Unidad de Informática para la Biodiversidad (UNIBIO) para reportar la presencia de diversos problemas relacionados con la recepción de ejemplares e imágenes usada.

La manera de utilizar la Interfaz de Reportes es la misma para todos los registros, sin importar de la colección que procedan.

En primer lugar, ingresar el número de folio o de catálogo del ejemplar a reportar en el campo “Folio”. Acto seguido, capturar en el campo “Fuente” el URL de la imagen. Finalmente, en el campo “Descripción” escribir el tipo de problema que se presenta, según sea el caso. Adicionalmente se incluye un campo de selección (Capturado) para indicar si el ejemplar se capturó en la Interfaz o no. Una vez terminado el reporte debe ser guardado, con ello la ventana se cierra y nuevamente se muestra la Interfaz de captura.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Es importante señalar que para cada problema en la captura de un registro, es necesario generar un reporte, de modo que, si un ejemplar presentase tres problemas, deben generarse tres reportes para ese registro.

### 4.3 Notas sobre la Interfaz de captura

A pesar de que para cada registro se despliegan los mismos campos, no todos los campos pueden ser completados con la información que cada ejemplar contiene. Por ejemplo, algunas etiquetas de ejemplares no poseen información referente a los campos de "Método de colecta" o "Profundidad"; incluso pueden no contener datos de campos más importantes como nombre científico o país. Es por ello que algunos campos pueden quedar vacíos y no es preciso escribir alguna nota o reporte para indicar la ausencia de los datos.

La opción de "Repetir anterior" se debe utilizar de manera adecuada y con criterio, por lo que es necesario corroborar que se respeta la captura verbatim, ya que como su nombre lo indica, repite íntegramente la información del ejemplar anterior en la sección en la que se use.

Debido a la heterogeneidad de las etiquetas de los ejemplares y con el fin de mantener los valores verbatim, es necesaria la creación de una serie de reglas de captura, que van desde lo general hasta lo particular, las cuales serán tratadas en la sección número siete de este manual.

## 5. Recepción de ejemplares e imágenes

Al recibir los ejemplares o fotografías para la captura de los datos es necesario verificar las condiciones en que se encuentran, pues deben cumplir con ciertas normas para ser capturados.

Primeramente, se debe revisar el estado del ejemplar, ya sean en ejemplares físicos o en imágenes de éstos. Para ambos, se ha de considerar que estén en óptimas condiciones, lo que garantiza que sea material de calidad y pueda ser publicado en el Portal de Datos Abiertos UNAM, Colecciones Universitarias.

En el caso de las imágenes de ejemplares, éstas deben ser nítidas, a fin de que las etiquetas asociadas a los ejemplares sean legibles. Además, debe existir congruencia de los datos de las etiquetas para extraer la información hacia la Interfaz de captura.

Para poder distinguir un ejemplar de otro, se emplea un número de folio o de catálogo. Es decir, un número de identificación único para cada ejemplar dentro de las distintas colecciones (folio en el caso de MEXU y número de catálogo para las colecciones de zoología), de tal manera que cada ejemplar debe contar con uno.

Aunque generalmente hay un ejemplar asociado a un número de identificación único, en casos específicos se puede presentar más de un individuo de la misma especie vinculado a un solo registro, por ejemplo en colecciones zoológicas como la Colección Nacional de Arácnidos (CNAN), la Colección Nacional de Crustáceos (CNCR) o la



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Colección Nacional de Moluscos (CNMO); pero todos pertenecen al mismo evento de colecta y están asociados al mismo número de identificación.

Existen otros montajes donde se encuentra más de un ejemplar de la misma o diferente especie, que corresponden a distintos eventos de colecta (p. ej., algunos ejemplares de la colección de algas de MEXU). Sin embargo, si es el caso, cada evento de colecta debe estar asociado a un número de identificación distinto. De no ser así, estos ejemplares deben ser reportados por medio de la Interfaz de Reportes de problemas en ejemplares.

Finalmente, los ejemplares de MEXU deben presentar el sello de la UNIBIO (Figura 3, inciso a) el cual tiene la función de indicar que dicho ejemplar ya fue fotografiado y puede continuar con el proceso de captura de datos. Otras colecciones, como las zoológicas, no presentan este sello, pero es debido a la naturaleza de éstas, por lo que no es considerado un error.

Existe otra serie de sellos, etiquetas y otros elementos que deben contener los ejemplares, pero eso dependerá del tipo de colección del que se trate. Por ejemplo, en el caso de MEXU todos los ejemplares deben tener el sello del herbario, que generalmente es de color rojo (Figura 3, inciso b), y en colecciones zoológicas un código QR (Figura 3, inciso c).



Figura 3. Sellos presentes en ejemplares. a) Sello de la UNIBIO y sus variantes. Se utiliza para indicar que el ejemplar ya ha sido fotografiado y continúa el proceso de captura de datos; b) Sello del Herbario Nacional de México (IBUNAM:MEXU:903262); c) Código QR asociado a registro de coleópteros de la Colección Nacional de Insectos (IBUNAM:CNIN:CO25456).

En el caso particular de MEXU, existen montajes que son fotografías de un ejemplar, es decir, no es un herborizado (Figura 4). Estas imágenes no deben ser ingresadas en la Interfaz de captura y se usa la Interfaz de Reportes de problemas en ejemplares para informar sobre ello.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 4. Fotografías de ejemplares. a) Los datos son ilegibles (IBUNAM:MEXU:334915); b) Datos legibles (IBUNAM:MEXU:230934).

### 5.1 Problemas en la recepción de ejemplares e imágenes

A continuación se describen los problemas registrados de las imágenes digitales de los ejemplares de algunas de las colecciones del IBUNAM.

La presencia de estos problemas se debe informar a través de la Interfaz de Reportes de problemas en ejemplares.

El primer aspecto a considerar dentro de los problemas con las imágenes de captura, es la ausencia de folio o número de catálogo, así como de los sellos de UNIBIO y los pertinentes para cada colección. De igual forma, tachaduras o enmendaduras en los números de folio, e incluso su legibilidad son motivo para generar un reporte. En casos especiales puede ocurrir que se presentan dos folios cuando sólo existe un evento de colecta, y no es posible asignar un folio único para el registro (Figura 5).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 5. Ejemplares que presentan inconvenientes que dificultan o imposibilitan la captura de los datos, por lo que se deben reportar por medio de la Interfaz de Reportes de problemas en ejemplares. a) No tiene sello de la UNIBIO ni folio; b) Carece de folio; c) Doble folio, uno con enmendadura y otro entre paréntesis (folio capturado IBUNAM:MEXU:120455); d) 321 no es un número de folio, por lo tanto el ejemplar carece de folio; e) El folio correcto es 6821 (IBUNAM:MEXU:6821), el número 2537 del sello color azul es una antigua numeración no vigente, también aparece el número 339 con formato similar al inciso d; f) Los números de catálogo entre el código de barras y QR son inconsistentes (IBUNAM:CNIN:HE4490 e IBUNAM:CNIN:HE4491).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

También deben ser considerados los problemas relacionados con el montaje de los ejemplares, es decir, más de un ejemplar de la misma o diferente especie. Sin embargo, algunos montajes de este tipo contienen datos aptos para su captura (Figura 6), pues son legibles y consistentes. Por lo que se debe proceder con la captura de estos ejemplares, según las reglas que se explican en la sección 7 de esta guía.

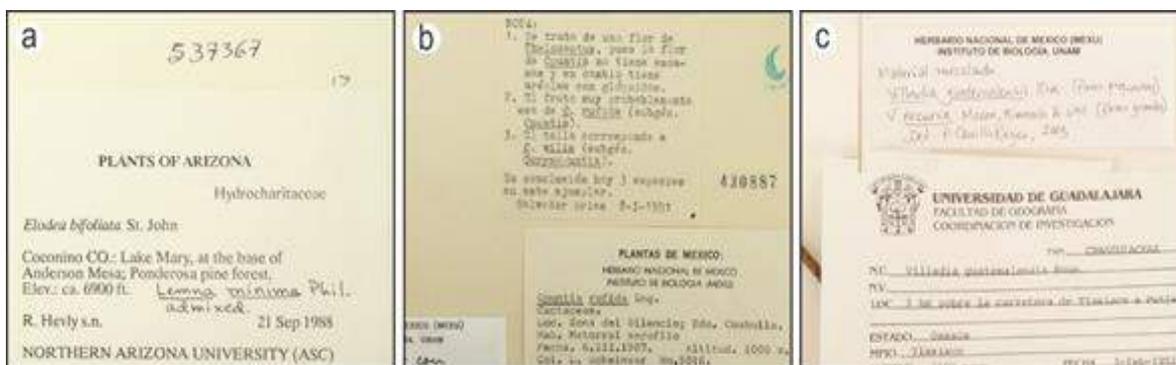


Figura 6. Montajes que contienen más de un ejemplar. Estos ejemplares deben reportarse en la Interfaz de Reportes de problemas en ejemplares. a) Dos ejemplares de diferente especie (IBUNAM:MEXU:537367); b) Tres ejemplares de diferente especie y cuál es la característica que las distingue (IBUNAM:MEXU:430887); c) Dos ejemplares de diferente especie, con una etiqueta de determinación referente (IBUNAM:MEXU:1157927).

Otros problemas están relacionados con la nitidez y legibilidad de las imágenes de los ejemplares. Se presentan generalmente como imágenes borrosas que fueron mal capturadas o dañadas durante la migración (Figura 7). En algunos casos, si parte de la información es legible se debe capturar como corresponda según las reglas de la sección 7 de esta guía.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 7. Imágenes borrosas. a) Imagen dañada; b) Mal enfocada al ser fotografiado el ejemplar (IBUNAM:CNIN:HE2669); c) Mal enfocada al ser fotografiado el ejemplar (IBUNAM:MEXU:562487).

### 6. Ingreso de registros

Una vez en la pantalla principal de ingreso a la Interfaz de captura, es necesario seleccionar la colección sobre la cual se hace el ingreso de datos y posteriormente el nombre del usuario y contraseña. Posteriormente, se despliega otra vista de la Interfaz, que aparece bloqueada a la escritura hasta que se ingrese el número de folio o número de catálogo.

Para comenzar la captura se verifica el folio del ejemplar en el campo del mismo nombre. Al realizar esta acción pueden ocurrir dos situaciones: la primera es que el ejemplar a capturar no se encuentre registrado por lo que se habilitarán todos los campos para proceder con la captura; y la segunda es que ya haya sido capturado, en tal caso debe informarse al encargado de la UNIBIO para decir el tratamiento del ejemplar.

Si el ejemplar es nuevo, después de habilitada la Interfaz, se procederá con la captura de todos los datos, más adelante se detalla la manera de realizar esta acción. En

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

el caso de las capturas realizadas por medio de imágenes digitales, el URL de la imagen debe ser capturado en el campo “Imagen de captura”.

Una vez ingresado cualquier valor en alguno de los campos aparece, en color rojo, la leyenda “Hay datos sin guardar” y ha de seguir así hasta que se guarde el registro. Para guardar el registro se emplea el botón “Guardar datos”. Una vez guardado, la leyenda desaparece y puede presionar el botón “Nuevo Registro”, lo que reinicia la Interfaz y queda bloqueada nuevamente para ingresar un folio nuevo. En la Figura 8 se puede apreciar parte de esta secuencia de actividades.

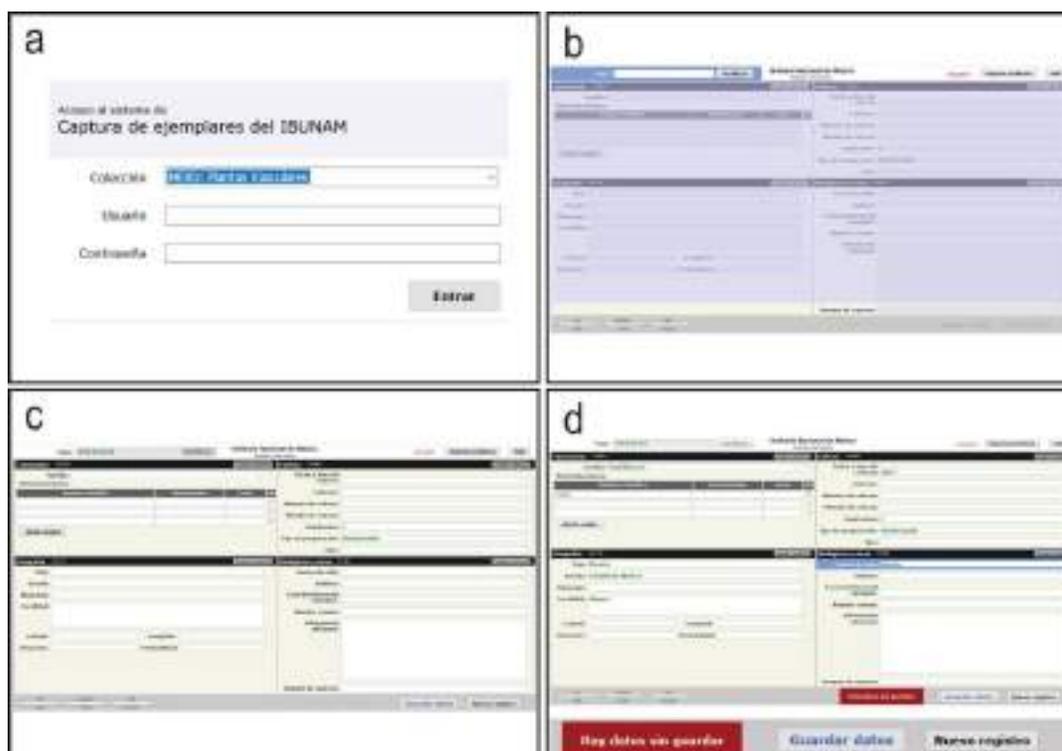


Figura 8. Interfaz de captura de la Unidad de Informática para la Biodiversidad (UNIBIO). a) Ventana principal de ingreso a la Interfaz de captura; b) Interfaz de captura bloqueada antes de ingresar el número de folio para verificar; c) Campos de la Interfaz de captura desbloqueados; d) Acercamiento de la sección sobre el control de captura.

### 7. Reglas de captura

En esta sección se establecen las reglas de captura para la información contenida en las etiquetas de los ejemplares de las colecciones biológicas del IBUNAM. Se encuentran agrupadas en dos secciones principales: generales y específicas. La descripción de las reglas de captura están acompañadas por figuras en las que se ejemplifican los casos

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

descritos en cada una. Algunas de estas figuras contienen un ejemplo de la manera correcta de cómo se debe de realizar la captura en la Interfaz.

Las reglas de captura, si bien corresponden a situaciones específicas, pueden ser acumulables en caso de que los datos de la etiqueta lo requieran a fin de mantener la información íntegra y literal de las etiquetas.

Dentro de las reglas de captura se emplean una serie de comandos que indican condiciones específicas, el modo y situaciones en las que se usan se detallarán más adelante. Los tres comandos principales son:

- de llaves (tecla F1; ingresa un triplete de llaves de apertura y de cierre, dentro del cual se ha de escribir la partícula en cuestión, {{{}}}; p. ej., {{{México}}}).
- de corchetes (tecla F9; ingresa un triplete de corchetes de apertura y de cierre, dentro del cual se ha de escribir la partícula en cuestión, [[[]]]; p. ej., [[[Flora Mexicana]]]).
- de diagonales o separador (tecla F4; ingresa un triplete de diagonales, y un espacio simple antes y después del triplete, /// ; p. ej., Abundante /// Floración: enero).

### 7.1. Reglas generales

Las reglas aquí incluidas aplican para todos los campos que contiene la Interfaz de captura, es decir, no responden a un tipo de dato en específico, sino que describen una situación en general.

#### 7.1.1 Captura total

Se debe capturar toda la información presente en el montaje del ejemplar, tanto aquella que se encuentra dentro como fuera de las etiquetas, esto incluye los datos escritos a mano. Los datos se capturan en su campo correspondiente o en el campo “Información adicional”, según sea el caso. Los sellos de MEXU y UNIBIO son la única excepción a esta regla (Figura 3).

##### 7.1.1.1 Diversidad de etiquetas

En este apartado se describirán los tipos de etiquetas presentes en los ejemplares de algunas de las colecciones del IBUNAM. Se distinguen dos tipos de etiquetas según la información que contienen.

###### 7.1.1.1.1 Etiqueta principal

Etiqueta que contiene los datos primarios, es decir, los datos que describen el evento de colecta del ejemplar en cuestión. En ejemplares de herbario, esta etiqueta generalmente se ubica en la esquina inferior derecha de la cartulina y contiene al menos los siguientes datos: colector, fecha y número de colecta, sitio de colecta y nombre científico (Figura 9).



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

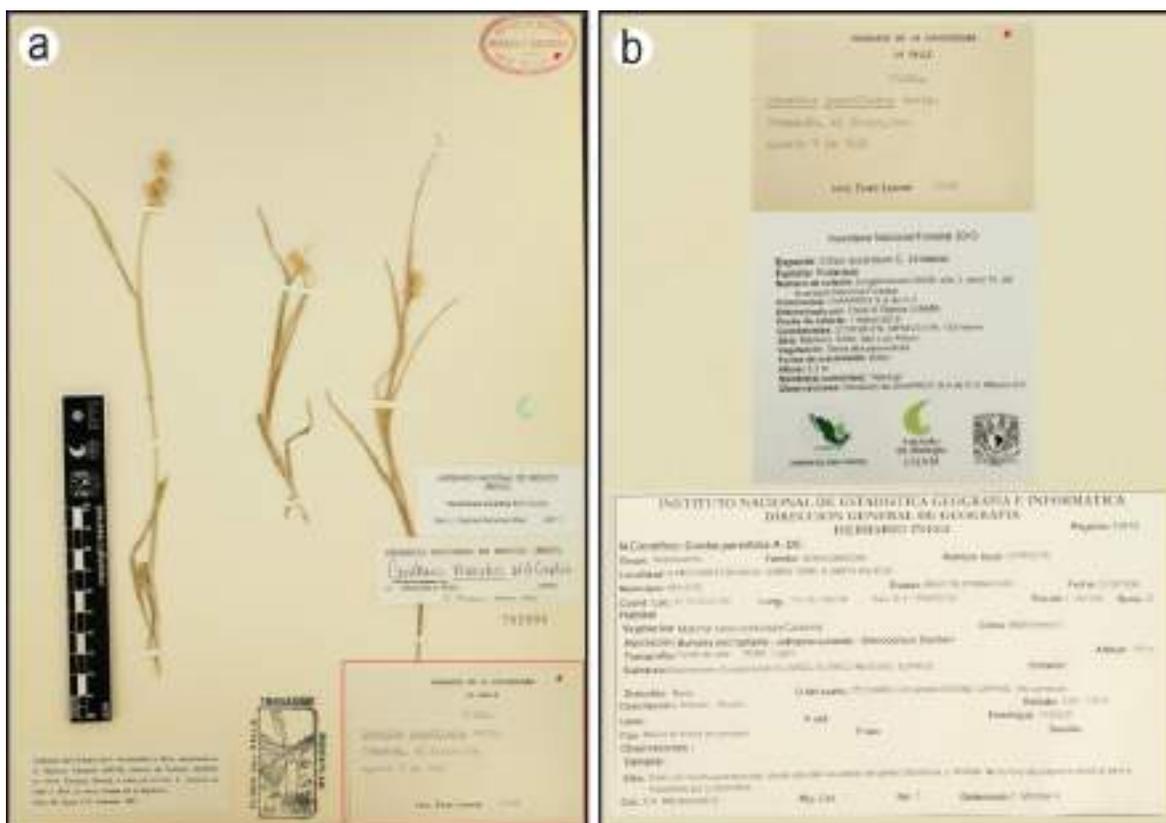


Figura 9. Etiqueta principal. a) Ubicación de etiqueta principal dentro el ejemplar de herbario. Se indica dentro de un recuadro rojo, para diferenciarse del resto de las etiquetas (IBUNAM:MEXU:282004); b) Ejemplos de diferentes tipos de etiquetas principales. Puede apreciarse la diferencia del número y tipo de dato que contienen, pero se mantienen los más importantes del evento de colecta (IBUNAM:MEXU:282004, IBUNAM:MEXU:1386351 y IBUNAM:MEXU:1405445).

### 7.1.1.1.2 Etiquetas secundarias

Éstas son anexas a la que contiene los datos primarios (etiqueta principal), pueden contener datos de diversa índole, como son: de añadidura, de corrección, de manejo curatorial, de toma de muestra, de notas taxonómicas, de determinación, etc.; como se describen a continuación:

#### 7.1.1.1.2.1 Adicional al evento

Puede tratarse de etiquetas que contengan datos primarios o de ocurrencia que se han añadido con la finalidad de completar o corregir un dato, en un tiempo posterior a la colecta e ingreso a la colección (p. ej., coordenadas, aclaración de la localidad o

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

municipio) (Figura 10). Se debe capturar el total de etiquetas y asignar los datos contenidos en ellas en los campos de la Interfaz de captura a los que correspondan según su naturaleza. Adicionalmente, la etiqueta debe ser capturada en su totalidad, sin atomizar (ver. “Atomización”), en el campo de “Información adicional”, con el uso del comando de corchetes, aun cuando toda la información de las etiquetas haya sido asignada a los campos correspondientes de la Interfaz.



Figura 10. Etiquetas adicionales al evento. Estas etiquetas contienen información sobre el evento de colecta, que no aparece en la etiqueta principal, es decir que complementan la información del evento de colecta. a) Contiene las coordenadas (IBUNAM:MEXU:1315117); b) Indica la localidad de colecta (IBUNAM:MEXU:1100832); c) Contiene dos pares de coordenadas y referencia a otra base de datos, así como el nombre de quien ingresó dicha información (IBUNAM:MEXU:388325).

### 7.1.1.1.2.2 Nota accesoria

Son las etiquetas que de manera exclusiva contienen información sobre el ejemplar o su manejo curatorial, sin contener datos primarios (Figura 11). Estos datos se deben capturar, en el campo “Información adicional”, completos y sin atomizar (ver. “Atomización”) con el uso del comando de corchetes. En caso de ser necesario, se pueden utilizar otras reglas de captura de forma acumulativa (ver “Captura verbatim”).

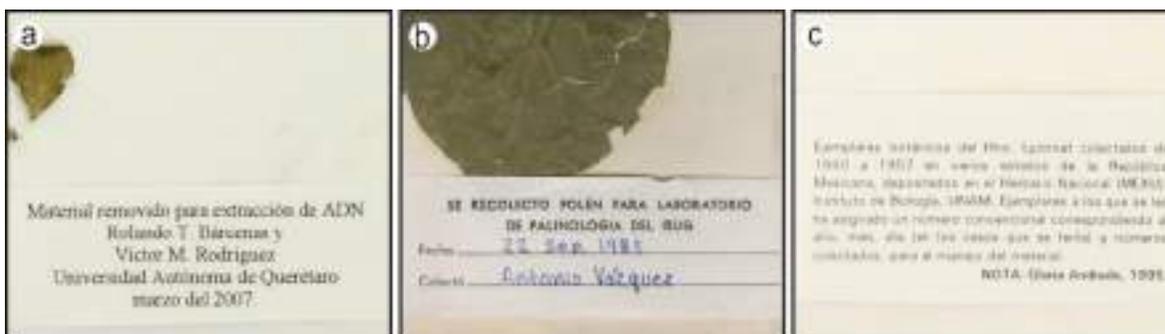


Figura 11. Etiquetas de Nota accesoria. a) Etiqueta de material removido (IBUNAM:MEXU:1047203); b) Etiqueta de toma de muestra de polen. Los textos de “Fecha” y

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

“Colectó” deben capturarse con el comando de llaves, ver “Captura verbatim”) (IBUNAM:MEXU:T20562); c) Etiqueta asociada al manejo curatorial (IBUNAM:MEXU:636727).

### 7.1.1.1.2.3 Texto asociado

Algunos ejemplares presentan etiquetas con información que corresponde a notas largas (p. ej., notas de difusión, divulgación, descripción, etc.) que no tienen correspondencia a algún campo de la Interfaz de captura y se excede la cantidad de información en etiquetas adicionales al evento (Figura 12). En estos casos se debe capturar en el campo “Información adicional” la leyenda: “Nota con texto asociado”, junto con el uso del comando de corchetes y no capturar la información de dicho texto asociado.



Figura 12. Etiquetas con textos asociados, en estos casos se excede la cantidad de información de etiquetas adicionales al evento o de nota accesoria. a) Hace referencia a un viaje de colecta y posee información de varios ejemplares colectados en el viaje, no sólo del montado (IBUNAM:MEXU:385711); b) Se muestra un artículo de una revista (IBUNAM:MEXU:120062).

### 7.1.1.1.2.4 Evento de determinación

Son etiquetas en las que se especifican los cambios de nombre científico resultado de las revisiones taxonómicas. El nombre científico puede o no variar a lo largo de los distintos eventos de determinación (Figura 13).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 13. Etiquetas de determinación. a) IBUNAM:MEXU:4868; b) IBUNAM:MEXU:1051866; c) IBUNAM:MEXU:126916; d) IBUNAM:MEXU: 1124414.

### 7.1.2 Asignación y atomización de la información

La información de las etiquetas debe ser ingresada en la Interfaz de captura, de modo que los datos se asocien a sus respectivos campos, siguiendo las reglas de asignación y atomización que se describen a continuación.

#### 7.1.2.1 Asignación

El proceso de asignación puede llevarse a cabo por medio de la identificación de los denominadores de campo, lo cuales permiten reconocer el tipo de dato dentro de las etiquetas.

Sin embargo, la asignación también puede realizarse por medio de la asociación semántica de la información cuando las etiquetas carecen de denominadores de campo. Este tipo de asociación puede realizarse sin dificultades, siempre que la información sea clara para asignar cada tipo de dato a su campo correspondiente, pero depende en gran medida del formato de la etiqueta.



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Si no es clara la asignación de un dato, ya sea porque su denominador de campo es ambiguo o carece de éste, dicha información debe capturarse en el campo de "Información adicional".

Debe evitarse separar de manera arbitraria información que forma parte de una oración dentro de la etiqueta si tal acción hace que pierda su significado original. Por lo que es recomendable leer una oración antes de comenzar la asignación. Este tipo de problemáticas son frecuentes en etiquetas con información asociada. Esto puede ocurrir con información tanto de la etiqueta principal como secundarias.

### 7.1.2.1.1 Asignación de información asociada

Existe información asociada a datos que corresponden a un determinado campo dentro de la Interfaz y que en primera instancia podría ser asignada a un campo diferente debido a su naturaleza, a pesar de que ambas partes tengan lógica por sí mismas. Sin embargo, al ser asignada por separado, esta información podría perder su sentido original causando interpretaciones erróneas en procesos posteriores a la captura.

Por lo tanto, en estos casos la información asociada se debe capturar en el campo correspondiente del tipo de dato principal al que están asociados. Por ejemplo en la Figura 14 se muestra información asociada al nombre científico, y que se debe capturar junto a éste en su campo correspondiente.

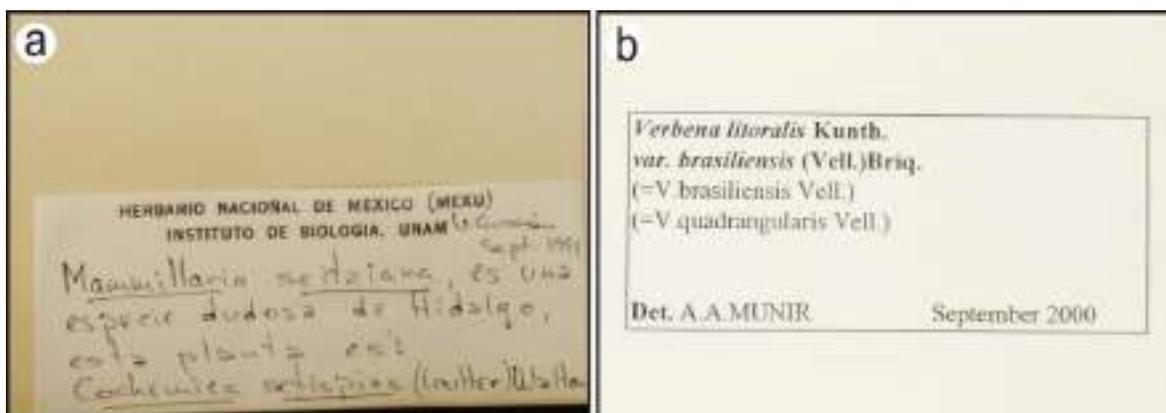


Figura 14. Etiquetas con información asociada. a) Contiene un comentario sobre la distribución de la especie propuesta para el ejemplar en cuestión y propone su determinación taxonómica (IBUNAM:MEXU:126911); b) Contiene sinonimias para el nombre científico de la determinación (IBUNAM:MEXU:521225).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.1.2.2 Atomización

Para el campo de “Información adicional” se debe realizar la atomización de la información, adicional al proceso de asignación.

La atomización consiste en separar las partículas (datos) que se asignan para capturar en dicho campo, esto facilita su procesamiento en etapas posteriores a la captura. Para separar las partículas asignadas se debe emplear el comando de diagonales o separador, el cual se coloca entre las partículas a atomizar.

La atomización de datos es exclusiva del campo “Información adicional”. La presencia de separadores en el resto de los campos es sólo para indicar doble dato (ver “Doble dato”), correcciones ortotipográficas (ver “Ortotipografía y correcciones”) o indicar el nombre de una familia asociada a una determinación anterior (ver “Varias Familias”); por lo que nunca se usa para atomizar los datos propios del resto de los campos, ya que para ello se realiza la asignación a los campos correspondientes.

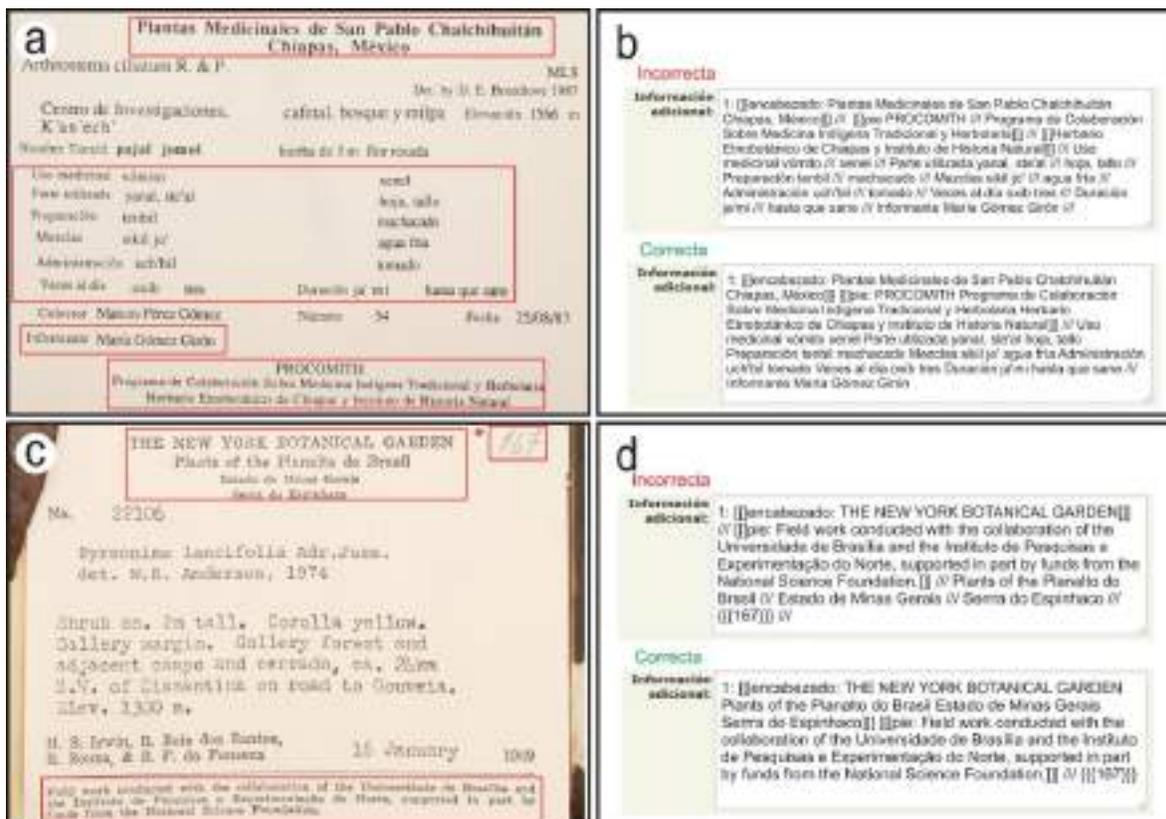
Así como ocurre con la asignación de información hacia los campos de la Interfaz, el uso del comando de diagonales (separador) en información adicional no debe ser usado excesivamente. La sobreatomización genera que se descontextualicen los datos, se pierda el orden y el sentido de la oración original, a pesar de que las partículas atomizadas estén capturadas en el mismo campo y una detrás de la otra. Esta problemática puede verse reflejada en un procesamiento erróneo de la información, ya que en el campo de “Información adicional” se capturan datos que si bien, no tienen un campo dentro de la Interfaz de captura, si lo tienen en el estándar Darwin Core 2; por tanto, la mala atomización genera un error en la asignación hacia sus campos correspondientes después de la captura.

Es importante mencionar que al ingresar el comando de diagonales se agregan automáticamente dos espacios, uno después de la última partícula escrita y otro antes de la que se comienza a escribir (partícula + espacio + /// + espacio + partícula), estos espacios deben ser respetados y evitar agregar un espacio al final de cada partícula, ya que el comando de diagonales lo hace automáticamente al ingresarse. Además, no se debe ingresar un separador al final del último dato capturado.

Los siguientes son ejemplos del uso adecuado y del abuso del comando de diagonales o separador (Figura 15). Las reglas específicas para la manera de ingresar los tipos de datos que se atomizan en el campo de “Información adicional” son explicadas, a detalle, en la sección 7.2 de esta guía.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



**a** **Plantas Medicinales de San Pablo Chulchihuitán Chiapas, México**  
*Arthrocnemum cilatum* K. & P. MEX  
 Det. by H. E. Robinson 1987  
 Centro de Investigaciones, capital, bosque y riego Elevación 1566 m  
 K'an'ech  
 Nombre local: pajal, jamael, kanta de San Francisco  
 Uso medicinal: sábilis, serol, hoja, tallo  
 Parte utilizada: jamael, raíz, raíz, tallo  
 Preparación: total machacado  
 Muestra: sábilis, raíz, agua fría  
 Administración: sábilis, machado  
 Fecha de recolección: 1987  
 Coleccionador: Martín Gómez Gil  
 Número: 34 Fecha: 25/05/87  
 Informante: María Gómez Gil  
 PROXIMITAD: Programa de Colaboración Sobre Medicina Indígena Tradicional y Herbolaria Herbario Etnobotánico de Chiapas y Instituto de Historia Natural

**b**  
**Incorrecta**  
 Información adicional: 1: [[encabezado: Plantas Medicinales de San Pablo Chulchihuitán Chiapas, México]] [[p: PROCOMITH]] [[Programa de Colaboración Sobre Medicina Indígena Tradicional y Herbolaria]] [[Herbario Etnobotánico de Chiapas y Instituto de Historia Natural]] [[Uso medicinal: sábilis]] [[Parte utilizada: jamael, raíz]] [[hoja, tallo]] [[Preparación: total]] [[machado]] [[Muestra: sábilis]] [[raíz]] [[agua fría]] [[Administración: sábilis]] [[Fecha de recolección: 1987]] [[Número: 34]] [[Fecha: 25/05/87]] [[Informante: María Gómez Gil]]

**Correcta**  
 Información adicional: 1: [[encabezado: Plantas Medicinales de San Pablo Chulchihuitán Chiapas, México]] [[p: PROCOMITH]] [[Programa de Colaboración Sobre Medicina Indígena Tradicional y Herbolaria]] [[Herbario Etnobotánico de Chiapas y Instituto de Historia Natural]] [[Uso medicinal: sábilis]] [[Parte utilizada: jamael, raíz]] [[hoja, tallo]] [[Preparación: total]] [[machado]] [[Muestra: sábilis]] [[raíz]] [[agua fría]] [[Administración: sábilis]] [[Fecha de recolección: 1987]] [[Número: 34]] [[Fecha: 25/05/87]] [[Informante: María Gómez Gil]]

**c** **THE NEW YORK BOTANICAL GARDEN**  
 Plants of the Planalto do Brasil  
 Estado de Minas Gerais  
 Serra do Espinhaço  
 No. 22106  
*Pyronia lanceifolia* Adr. Fum.  
 det. W. R. Anderson, 1974  
 Shrub ca. 2m tall. Corolla yellow.  
 Gallery margin. Gallery forest and adjacent campo and cerrado, ca. 2km S.W. of Consistência on road to Consistência.  
 Elev. 1330 m.  
 H. S. Gentry, H. Reis dos Santos, R. Soares, & R. F. de Faria 16 January 1969  
 Field work conducted with the collaboration of the Universidade de Brasília and the Instituto de Pesquisas e Experimentação do Norte, supported in part by funds from the National Science Foundation.

**d**  
**Incorrecta**  
 Información adicional: 1: [[encabezado: THE NEW YORK BOTANICAL GARDEN]] [[v]] [[p: Field work conducted with the collaboration of the Universidade de Brasília and the Instituto de Pesquisas e Experimentação do Norte, supported in part by funds from the National Science Foundation]] [[v]] [[167]] [[v]]

**Correcta**  
 Información adicional: 1: [[encabezado: THE NEW YORK BOTANICAL GARDEN]] [[v]] [[p: Field work conducted with the collaboration of the Universidade de Brasília and the Instituto de Pesquisas e Experimentação do Norte, supported in part by funds from the National Science Foundation]] [[v]] [[167]] [[v]]

Figura 15. Atomización de datos. Uso del comando de diagonales o separador en el campo de “Información adicional”. a) y b) Se muestra la etiqueta de colecta y en recuadros rojos los elementos que se deben capturar en el campo de “Información adicional”, así como un ejemplo de captura incorrecta y correcta de los datos (IBUNAM:MEXU:579582); c) y d) Muestran un ejemplo similar, pero con otra etiqueta de colecta (IBUNAM:MEXU:183865).

En general, tanto para la asignación y atomización se ha de considerar como parte de una buena práctica no separar el adjetivo del sustantivo que califica, y que las partículas por sí solas no son un dato obvio, principalmente cuando carecen de un denominador de campo.

### 7.1.2.3 Doble dato

Un caso especial, principalmente para la asignación, es el “doble dato”. Es decir, que si en la misma o diferente etiqueta se presentan dos datos de la misma naturaleza, distintos o repetidos, ya sea añadidura o cualquier otra situación, ambos datos deben ser capturados en el campo correspondiente y separados con el comando de diagonales.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

De ser posible, debe seguirse el orden cronológico en que aparecen, es decir, primero el dato de la etiqueta principal y después el añadido o procedente de una etiqueta secundaria. En el caso de ser datos dobles que deben ser asociados a dos o más campos relacionados seguir la relación lógica que presentan (Figura 16).



Figura 16. Doble dato. a) Ejemplar con doble dato de localidad (IBUNAM:MEXU:123398); b) Etiqueta con doble dato de coordenadas geográficas (IBUNAM:MEXU:1409797); c) Ejemplar con dato doble de colector y número de colecta (IBUNAM:MEXU:1265320).

Los casos que incluyen las correcciones o añadiduras escritas a mano, se detallarán más adelante. Sin embargo, los eventos de determinación presentes o no en la misma etiqueta, deben ser considerados como entradas diferentes.

Si algunos datos considerados como dobles deben ser capturados en el campo de "Información adicional" y dado que el separador genera dos partículas independientes, se debe capturar los dos datos asociados a las partículas (dato 1 y dato 2) sin emplear el comando de diagonales, únicamente separados ambos datos por un espacio sencillo (p. ej., [...] /// dato 1: florece de enero a marzo dato 2: florece a lo largo de todo el año /// [...]). Toda la cadena de texto que se genere, debe ser considerada como una sola partícula para fines de la atomización en este campo.

### 7.1.3 Captura verbatim

Las obligaciones y derechos que se tienen sobre el manejo de los datos en la cadena de digitalización es diferencial de acuerdo a la etapa que se trate. Por lo que algunos de los cambios a realizar no se encuentran dentro de los alcances del capturista, y únicamente les corresponden a los curadores. Así que, se debe evitar corregir o modificar los datos de los ejemplares al capturarlos, incluso aunque tengan errores evidentes.

Si la información no es clara, es necesario informar sobre el ejemplar a través de la Interfaz de Reportes de problemas en ejemplares, pero nunca se debe interpretar ni modificar la información.

De tal manera que con el objetivo de mantener los datos "verdaderos" e "íntegros" contenidos en las etiquetas, es necesario que se realice la captura de manera literal, palabra por palabra, incluso cuando se detecten faltas ortográficas, abreviaturas,

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

inconsistencias, etc. Para auxiliar al capturista en la captura verbatim, existe una Interfaz accesoria (“UNIBIO Web Caracteres”), en la que aparecen un conjunto de símbolos, caracteres poco comunes y algunas de sus variantes en el alfabeto latino, tales como acentos diacríticos (p. ej., à, â, ã, ä, etc.). Para más referencias sobre etiquetas escritas en idiomas distintos al español revisar los apartados “En dos idiomas” y “Extranjeras” en esta misma sección.

A continuación, se presentan reglas puntuales sobre la captura verbatim de tópicos específicos, los cuales se ejemplifican de manera gráfica.

### 7.1.2.1 Sin formato de control

#### 7.1.3.1.1 Error de origen

Cuando se ha detectado algún error en la etiqueta y se tiene una referencia confiable de los datos correctos. En estos casos es necesario preservar los datos originales sin importar el tipo de etiqueta que se trate, por tanto, la captura debe ser total y verbatim indistintamente del tipo de error encontrado (tipográfico, pertenencia, inconsistencia entre los datos, etc.) (Figura 17). Se debe informar de la situación a través de la Interfaz de Reportes de problemas en ejemplares.



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Figura 17. Etiquetas con error de origen. a) El apellido del segundo colector está mal escrito (IBUNAM:MEXU:419992); b) Contiene dos familias. En este caso se añade la regla de doble dato para mantener la captura verbatim (IBUNAM:MEXU:1387964); c) Fecha de colecta incompleta (IBUNAM:MEXU:269527); d) Sello de fecha de colecta con un dígito adicional (IBUNAM:MEXU:1402039).

### 7.1.3.1.2 Probable error

Puede aplicarse a los casos en los que no se tenga la certeza del error "detectado" en la etiqueta (Figura 18). Se debe realizar la captura verbatim e informar de la situación a través de la Interfaz de Reportes de problemas en ejemplares.

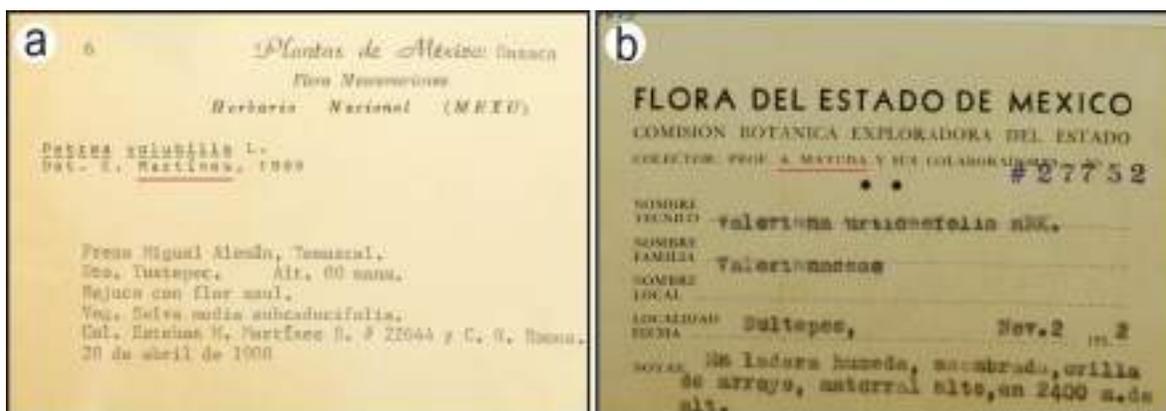


Figura 18. Etiqueta con probable error. a) Probablemente se trate de E. Martínez (IBUNAM:MEXU:635959); b) Probablemente se trate de E. Matuda (IBUNAM:MEXU:109875).

### 7.1.3.1.3 Variante

Se aplica a etiquetas que presentan diferencias en los datos cuando pertenecen a elementos concordantes. Por ejemplo, variaciones en el nombre de los colectores, diferentes años de identificación para la misma especie, etc. La captura debe ser verbatim para casos como estos (Figura 19).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Figura 19. Etiquetas con variantes en el nombre del colector. a) H. Asbjornson (IBUNAM:MEXU:1282877); b) H. Asbjorsen (IBUNAM:MEXU:1239927); c) H. Asborgsen (IBUNAM:MEXU:1318784).

### 7.1.3.1.4 Mayúsculas

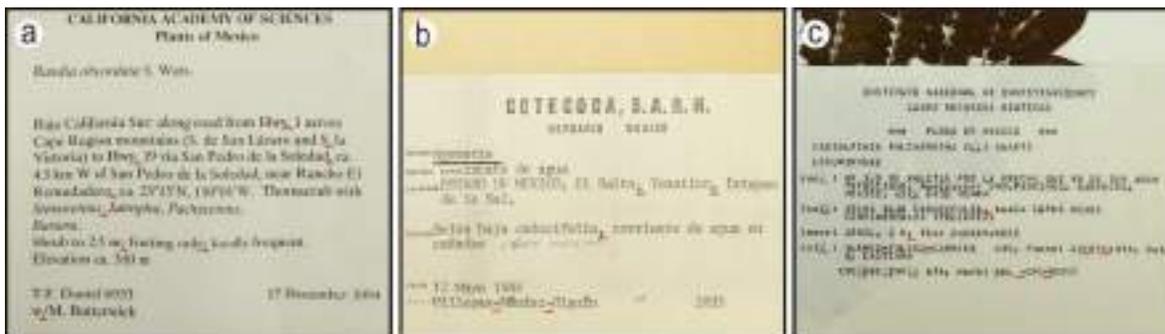
Dentro de los distintos tipos de etiquetas existen algunas que total o parcialmente están escritas con mayúsculas (Figura 20). Es preciso capturar de manera verbatim, sin importar si se trata de nombres científicos o propios, a pesar de ser técnica y ortográficamente incorrecto. Incluso si se combinan mayúsculas y minúsculas en la misma palabra u oración.



Figura 20. Etiquetas escritas con mayúsculas. a) Toda la información incluyendo el nombre científico están escritas en mayúsculas (IBUNAM:MEXU:263792); b) El encabezado y los nombres de los colectores aparecen en mayúsculas (IBUNAM:MEXU:144974).

### 7.1.3.1.5 Signos de puntuación

Se requiere respetar cualquier signo de puntuación o cualquier otro símbolo que contenga la etiqueta (Figura 21). Por lo general, éstos funcionan para separar los datos. La captura debe ser verbatim.



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Figura 21. Signos de puntuación utilizados como separadores de datos. a) Presentes principalmente en la descripción de la localidad (IBUNAM:MEXU:768830); b) Presentes en localidad, hábitat y colectores (IBUNAM:MEXU:1083378); c) Presentes en colectores y fecha de colecta, principalmente (IBUNAM:MEXU:843056).

Sin embargo, es igualmente importante respetar la ausencia de signos de puntuación, para obedecer la captura verbatim, aunque ortográficamente sea incorrecto o se intuya que los datos debieran estar separados con signos de puntuación (Figura 22). En caso de que la información esté separada por un salto de línea, se debe capturar un espacio sencillo, más nunca un salto de línea (tecla *enter*).

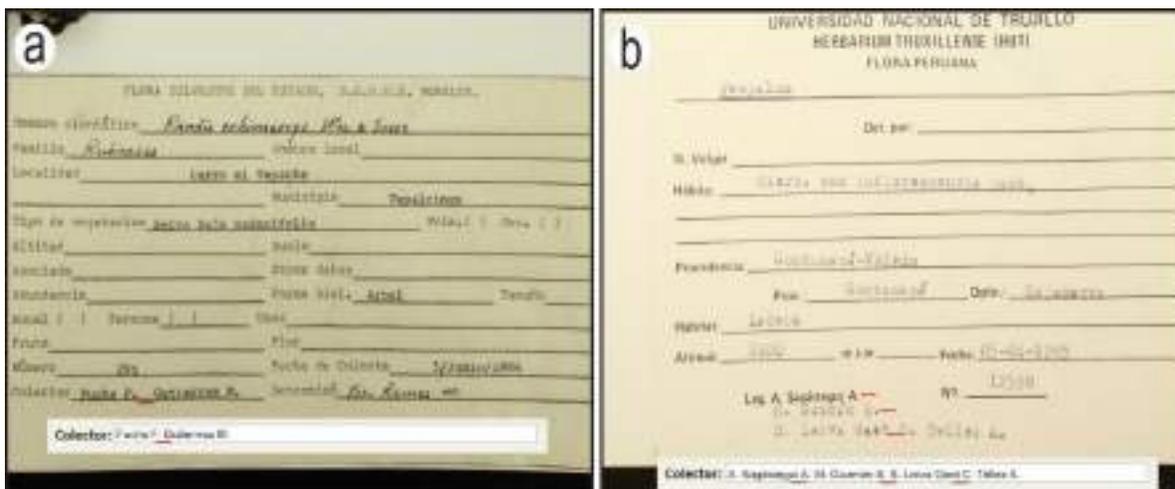


Figura 22. Etiquetas sin el uso de signos de puntuación como separadores. a) Los nombres de los colectores están separados por un espacio (IBUNAM:MEXU:790463); b) Los nombres de los colectores aparecen en renglones diferentes, pero no existen signos adicionales para separarlos (IBUNAM:MEXU:965378).

### 7.1.3.1.6 Denominador de campo

Los denominadores de campo permiten identificar el tipo de dato dentro de las etiquetas. Esto es útil para asignar correctamente los datos al campo que les corresponde dentro de la Interfaz de captura. Por lo que si el denominador de campo es el mismo que en la Interfaz de captura, éste no se debe capturar nuevamente en el campo que corresponda, sino solo el dato. Sin embargo, si el denominador de campo presenta alguna variante (p. ej., un sinónimo o abreviatura) éste debe ser capturado en el campo correspondiente.

Por otro lado, si la información de la etiqueta presenta denominadores de campo que no aparecen en la Interfaz de captura, se debe capturar el denominador de campo junto al dato que le corresponde en el campo “Información adicional”. Se debe evitar la omisión de los denominadores de campo porque la información podría perder su sentido original, a pesar de que el denominador de campo pudiera ser obvio o redundante (Figura 23).



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Cabe señalar que si existe un denominador de campo sin dato asociado, éste no debe ser capturado, sin importar si aparece o no en la Interfaz de captura.

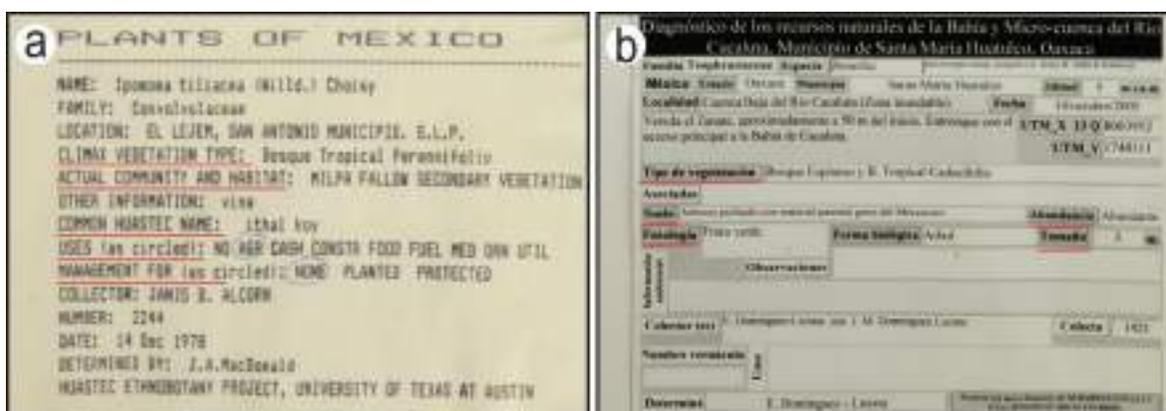


Figura 23. Denominador de campo. a) Denominadores de campo que deben capturarse para que no se pierda el contexto del dato (IBUNAM:MEXU:757902); b) Varios de los denominadores de campo no deben capturarse, pues tienen su correspondiente campo en la Interfaz de captura, pero otros presentan variaciones que deben capturarse asociados a los datos (IBUNAM:MEXU:1268227).

En el caso de etiquetas que carezcan de todos o algunos de los denominadores de campo y sólo contienen los datos, estos pueden ser asignados a los campos correspondientes dentro de la Interfaz, según el tipo de dato de que se trate. En caso de que no sea clara la identificación del tipo de dato se deben capturar los datos en el campo "Información adicional", para ello se usa el comando de diagonales para separar la información (Figura 24).

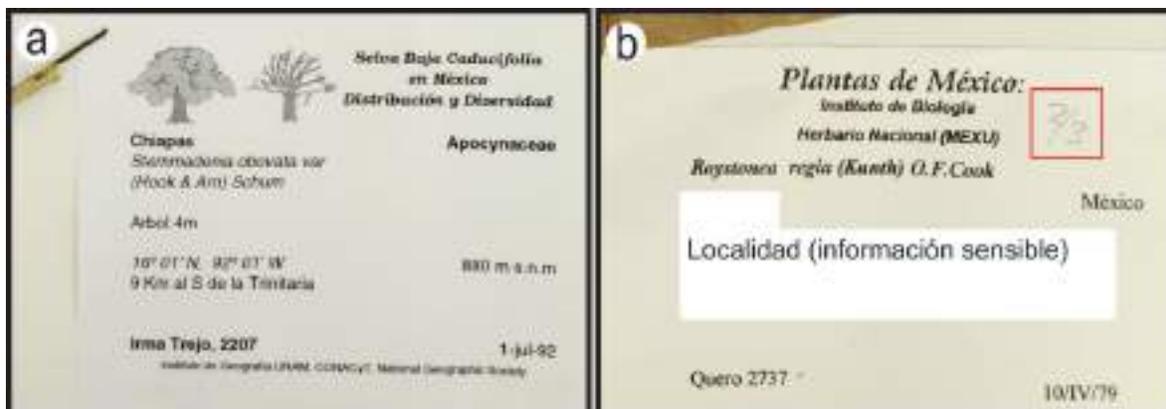


Figura 24. Etiquetas que carecen de denominadores de campo. a) Etiqueta sin denominadores de campo pero que probablemente no dificulte la captura (IBUNAM:MEXU:1100832); b) La falta del

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

denominador genera incertidumbre sobre la naturaleza del dato. Sin embargo, debe ser capturado en el campo de "Información adicional" (IBUNAM:MEXU:1006246).

### 7.1.3.1.7 En dos idiomas

Son etiquetas que contienen información en dos idiomas, puede tratarse de cualquier lengua distinta al español, ya sea extranjera u originaria de México (Figura 25). La captura debe ser verbatim, sin traducir ni omitir información; esto incluye los denominadores de campo en el idioma distinto al español dentro de la etiqueta, así como los que no aparecen en la Interfaz de captura y sus variantes.

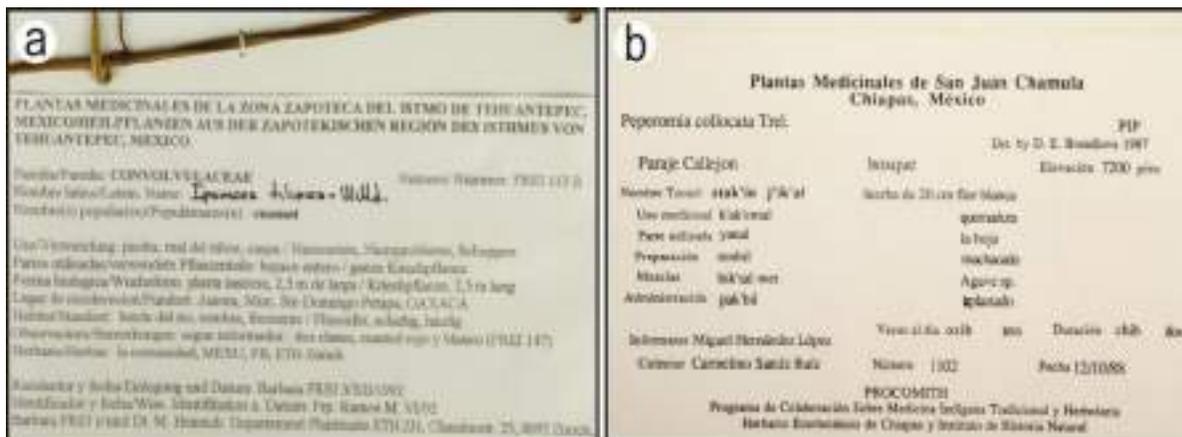


Figura 25. Etiquetas en dos idiomas. a) Alemán y español (IBUNAM:MEXU:713487); b) Español y tzotzil (IBUNAM:MEXU:572780).

### 7.1.3.2 Con formato de control

#### 7.1.3.2.1 Escrita a mano

Toda aquella información escrita a mano (pluma, lápiz o similar) debe ser capturada por medio del comando de llaves en su campo correspondiente. Según el caso, pueden presentarse distintas combinaciones de captura, como las que se muestran en la Figura 26. Es decir, todo el valor para un campo puede estar escrito completamente a mano o tener texto a máquina y adiciones hechas a mano. Esta regla aplica incluso a datos ilegibles (ver "Ilegible").

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

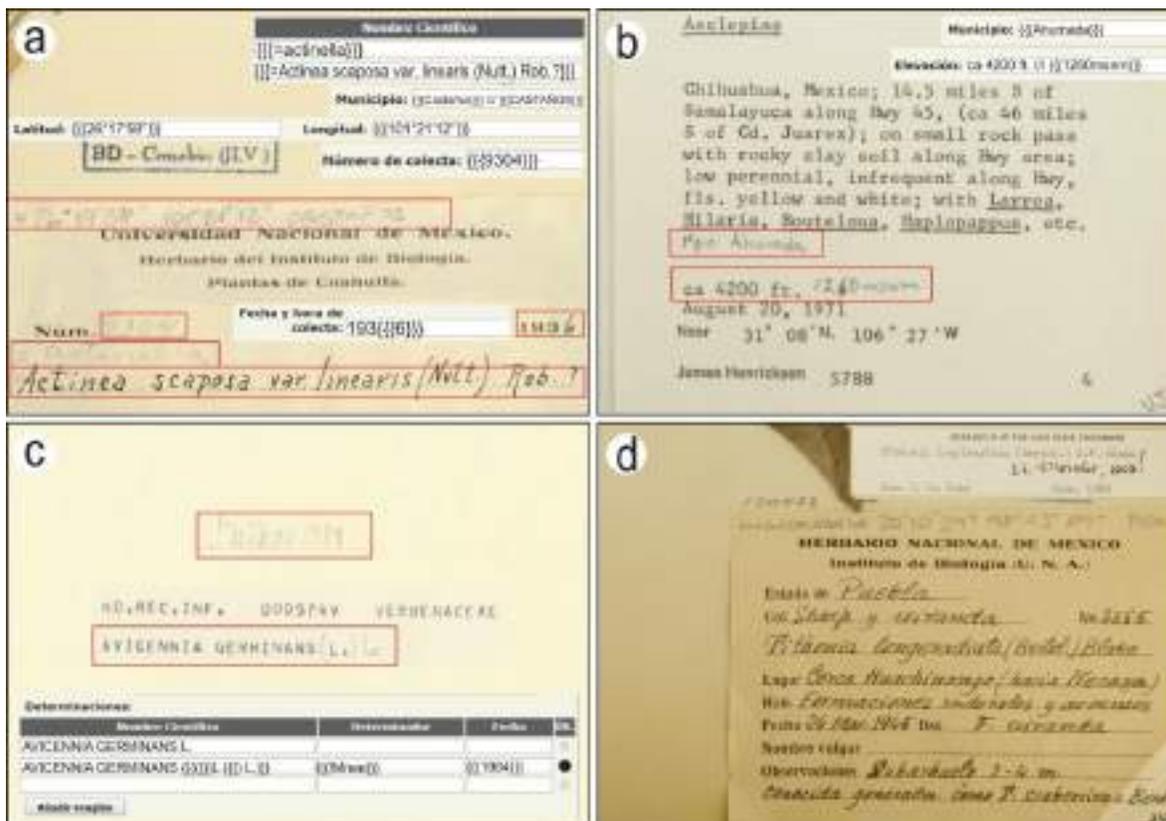


Figura 26. Etiquetas que presentan datos escritos a mano. Se muestra la forma correcta de capturar algunos de los datos escritos a mano como ejemplo. a) Las coordenadas y el nombre del municipio fueron añadidas a mano, así como un evento distinto de determinación. Igualmente el número de colecta y parte de la fecha están escritas a mano (IBUNAM:MEXU:33883); b) A manera de corrección fue añadida la altitud pero sin bloquear el primer valor, por lo que se captura como dato doble y se adiciona el valor de municipio (IBUNAM:MEXU:1045518); c) Última determinación (determinador y año) escrita a mano (IBUNAM:MEXU:173285); d) Toda la etiqueta está escrita a mano, con adiciones y última determinación (determinador y año) (IBUNAM:MEXU:120422).

### 7.1.3.2.2 Ilegible

Estos son datos ininteligibles que no se pueden capturar, ya sea porque la tinta de la etiqueta se encuentra en malas condiciones, o la escritura no es clara (Figura 27). Se debe capturar lo que sea legible y la leyenda “ilegible” a modo de sustitución de lo que no se entienda, que pueden ser desde una palabra hasta una oración entera. No se deben usar signos de interrogación en sustitución del dato ilegible (p. ej., colector: ilegible). Generalmente este tipo de datos están escritos a mano, por lo que el uso del comando de llaves es aplicable a este tipo de datos.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Esta leyenda puede ser incluida en el campo correspondiente del dato por el cual se sustituye. Sin embargo, si estos datos no pueden ser ubicados en algún campo en concreto (ya sea porque la ausencia de un denominador de campo o éste no aparece en la Interfaz), la leyenda “ilegible” se debe capturar en el campo “Información adicional”. Estos casos se informarán a través de la Interfaz de Reportes de problemas en ejemplares.



Figura 27. Etiquetas con contenido ilegible. a) Nombre científico donde el epíteto específico es ilegible (IBUNAM:MEXU:371169); b) El apellido del determinador es ilegible (IBUNAM:MEXU:80787); c) El nombre científico es ilegible (IBUNAM:MEXU:1327119); d) Etiqueta con la mayoría de los datos ilegibles. Aparecen referidos el nombre común, el nombre científico y la localidad (IBUNAM:MEXU:524396).

### 7.1.3.2.3 Extranjera

Aplica a las etiquetas que están escritas en otro idioma diferente al español. Estos ejemplares pueden ser originarios de otros países o provenir de colecciones extranjeras. Además, pueden estar escritas exclusivamente en algún alfabeto distinto al abecedario



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

latino (Figura 28). Para estos ejemplares pueden presentarse varias circunstancias, pero de manera general no está permitido realizar ningún tipo de traducción.

En caso de que el idioma extranjero esté basado en el alfabeto latino, se debe capturar la información en los campos correspondientes. Es recomendable el uso de la Interfaz accesoria “UNIBIO Web Caracteres” para ingresar símbolos o letras poco comunes, como letras con acentos diacríticos.

Por el contrario, si el idioma no está basado en el alfabeto latino, se debe capturar lo reconocible en los campos correspondientes. Lo que no pueda ser capturado ha de ser sustituido por la leyenda “otro alfabeto”, junto con el comando de corchetes (p. ej., [[[otro alfabeto]]]). Esta leyenda puede ser ingresada en cualquiera de los campos de la Interfaz, siempre y cuando exista un denominador de campo claro en la etiqueta para asignarlo a dicho campo. Si el denominador de campo no es preciso o no existe, la leyenda debe ser capturada en el campo de “Información adicional”.

Estos ejemplares se deben reportar en la Interfaz de Reportes de problemas en ejemplares, especialmente cuando la cantidad de información capturada sea mínima.

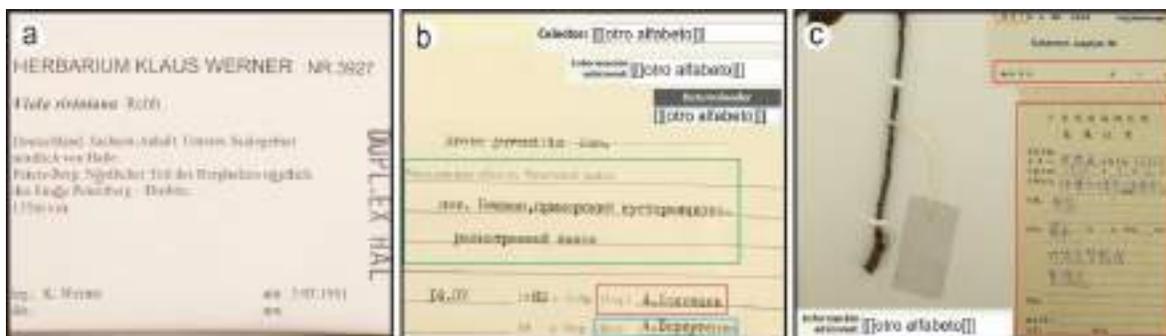


Figura 28. Etiquetas extranjeras. Se muestran algunos ejemplos de captura de la información. a) Escrita en alemán. La información puede ser capturada y debe mantenerse la puntuación de las palabras (IBUNAM:MEXU:1256632); b) Escrita en ruso, pero con algunos denominadores de campo en alfabeto latino (IBUNAM:MEXU:578759); c) La etiqueta principal está escrita con caracteres distintos al alfabeto latino, incluyendo denominadores de campo, así como algunas partes de la etiqueta de determinación (IBUNAM:MEXU:252509).

### 7.1.3.2.4 Confusa

Son aquellas etiquetas que presentan la información de manera imprecisa, desordenada o sin referencia y es difícil distinguir los datos certeros (Figura 29). Se debe capturar el dato junto con la leyenda “confuso” y con el comando de corchetes (p. ej., [[[dato confuso]])] e informar a través de la Interfaz de Reportes de problemas en ejemplares.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

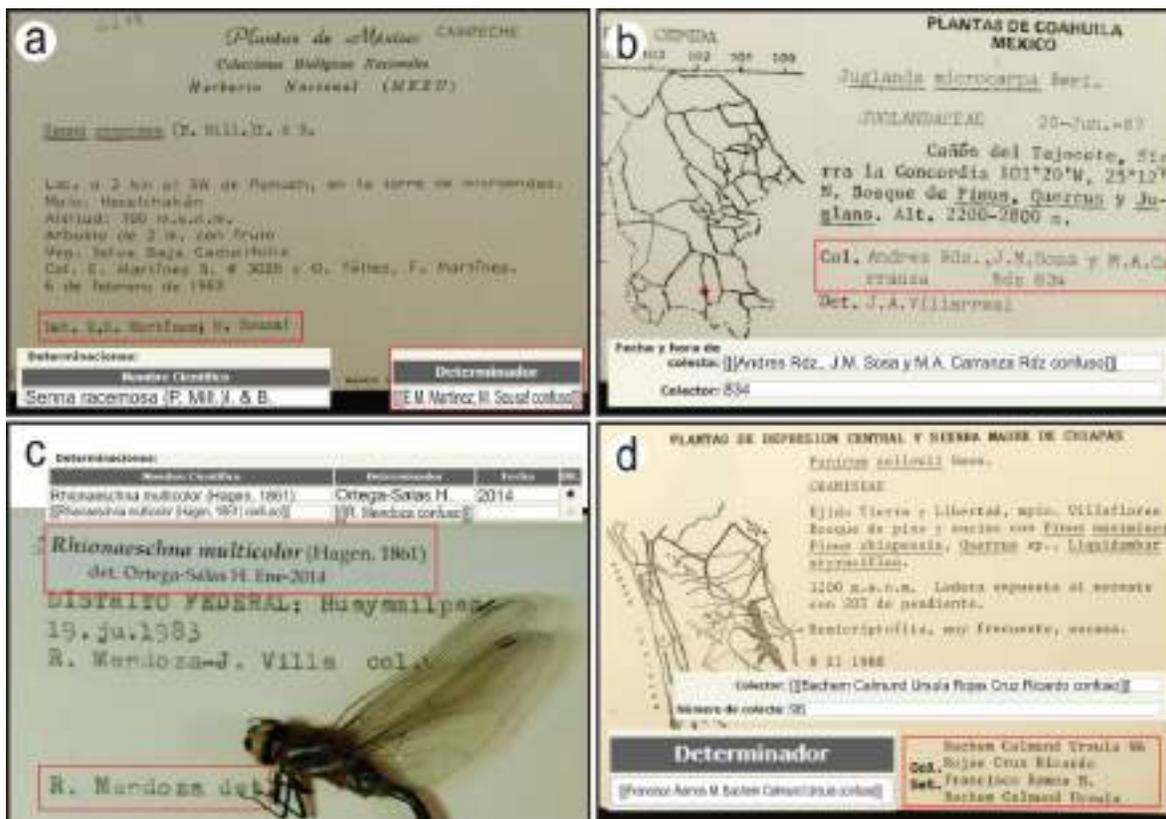


Figura 29. Etiquetas confusas. a) No son muy claros los primeros eventos de determinación (IBUNAM:MEXU:345996); b) Es dudosa la pertenencia de la partícula "Rdz" (IBUNAM:MEXU:472136); c) Existen dos determinadores sin coherencia entre los datos (IBUNAM:CNIN:OD7651); d) No es claro cuáles son los nombres que corresponden a los colectores y determinadores (IBUNAM:MEXU:895520).

### 7.1.3.2.5 Opciones múltiples, de llenado o marcado

Estas etiquetas presentan opciones de llenado para el registro de los datos. Sólo deben ser capturadas las opciones que contienen datos o estén marcadas; capturando el denominador de campo cuando no exista en la Interfaz de captura o presente alguna variante. En el caso de que no exista marcada ninguna opción a elegir, no se debe capturar ninguna de ellas y esta situación se reporta en la Interfaz de Reportes de problemas en ejemplares, y en los campos correspondientes de la Interfaz de captura se debe indicar con la leyenda "sin opcion" (debe escribirse sin acento en la segunda letra "o") con el uso del comando de corchetes (p. ej., [[sin opcion]]) (Figura 30).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 30. Etiquetas con opciones múltiples para el registro de los datos del ejemplar. a) Etiqueta con casillas para llenado (IBUNAM:MEXU:1004979); b) Etiqueta con datos seleccionadas con círculo (IBUNAM:MEXU:406356); c) Existen varias opciones en las que se incluyen la localidad, el hábitat, la elevación y la fecha de colecta, pero no se indica alguna de las opciones disponibles (IBUNAM:MEXU:170134).

### 7.1.3.2.6 Múltiples

Cuando se presentan múltiples etiquetas principales puede ocurrir que sean consistentes. Es decir, que cuando se presentan dos o más etiquetas asociadas al mismo ejemplar, puede tratarse de etiquetas históricas, nuevas o duplicadas (Figura 31). Se debe capturar la información total de todas las etiquetas, y extraer los datos para los campos de la Interfaz; se aplica la regla para “doble dato” en caso de ser necesario. Adicionalmente, la etiqueta antigua (o duplicada) se debe capturar en el campo de “Información adicional” sin atomizar como si fuera una etiqueta de nota accesoria, para ello se emplea el comando de corchetes (ver “Nota accesoria”).

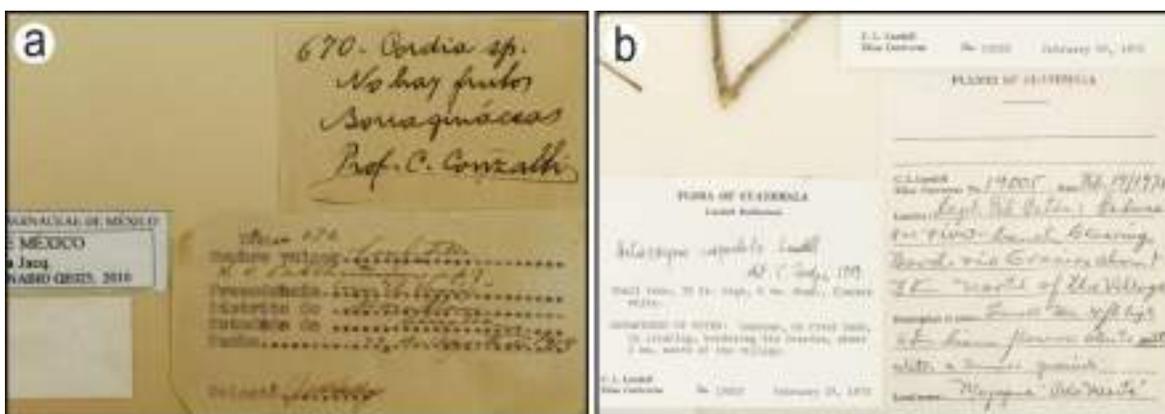


Figura 31. Ejemplares con múltiples etiquetas consistentes. a) La información de las etiquetas se complementa (IBUNAM:MEXU:169244); b) Ejemplar en el que una de las etiquetas se transcribió a partir de la otra (IBUNAM:MEXU:540476).

Sin embargo, puede darse el caso en el que las etiquetas múltiples sean inconsistentes. Es decir, que cuando se presentan dos o más etiquetas asociadas al



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

mismo ejemplar, estas pueden pertenecer a eventos de colecta diferentes; incluso si parecen etiquetas históricas, nuevas o duplicadas. Por lo que la información de estas etiquetas puede ser incongruente entre ellas (Figura 32).

En estos casos se debe capturar todos los datos de todas las etiquetas a los campos correspondientes de la Interfaz de captura; se aplica la regla de “doble dato” en caso de ser necesario. Así mismo, se debe capturar la etiqueta más antigua o duplicada inconsistente sin atomizar en el campo de “Información adicional”. De igual forma se debe usar la Interfaz de Reportes de problemas en ejemplares para informar sobre estos casos.

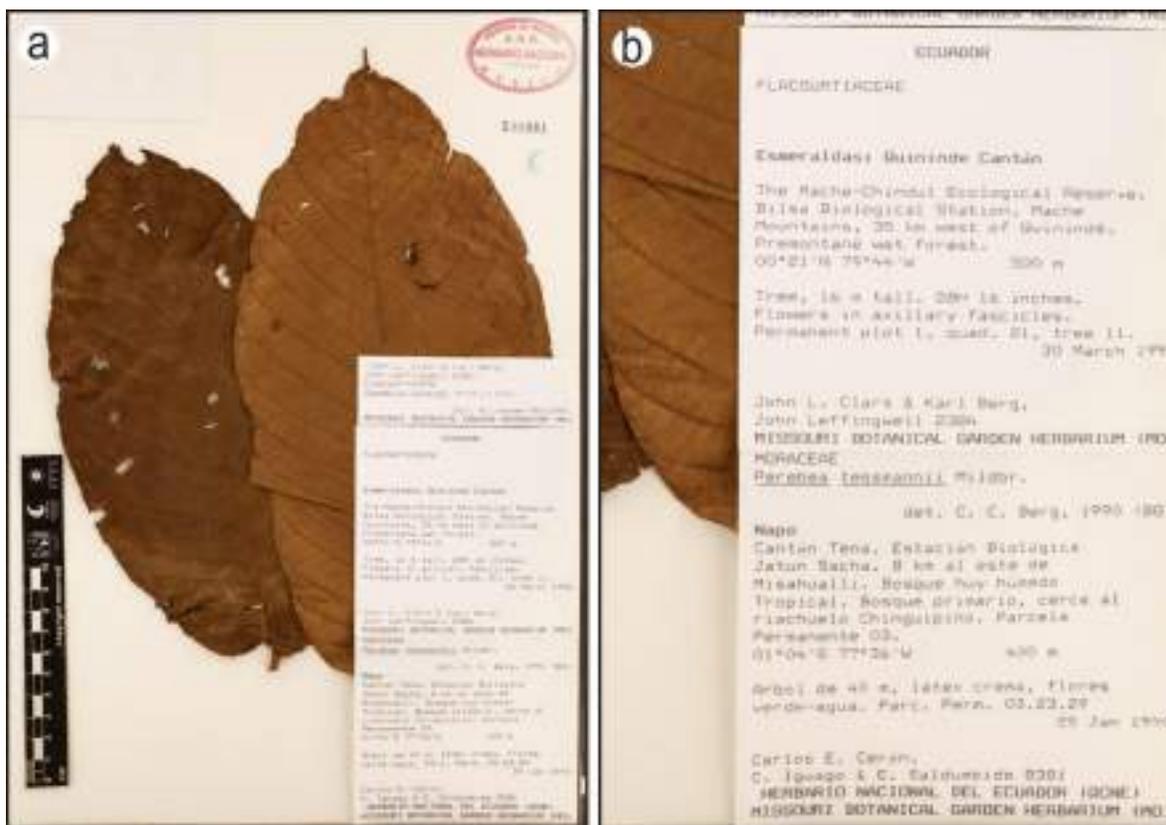


Figura 32. Ejemplares con etiquetas múltiples inconsistentes. a) Ejemplar con etiquetas múltiples de dos eventos de colecta diferentes (IBUNAM:MEXU:930981); b) Acercamiento de inciso “a” para visualizar las inconsistencias. Todos los datos son completamente diferentes entre las etiquetas.

### 7.1.3.2.7 Ortotipografía y correcciones

Cuando se presenten correcciones en la información de las etiquetas que involucren cambios ortotipográficos, se deben capturar los datos originales y la corrección sugerida en su campo correspondiente. Sin importar si la modificación es correcta o no. Se debe



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

Sugerencia de cita: DGRU.2020. Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2 (Versión 1). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

usar el comando de tripletes para separar ambos datos, y escribir primero el dato original y luego el dato corregido (Figura 33). En estos casos no se utilizan los comandos de llaves o de corchetes para indicar letras escritas a mano o tachadas (ver. “Enmendadura”), respectivamente.

Si la corrección ocurre dentro de una lista o serie de elementos, como un equipo de colecta, se debe considerar la lista completa como si fuese una partícula única y repetirse nuevamente después del triplete de diagonales indicando la corrección (Figura 33).

Si este tipo de correcciones son realizadas sobre nombres científicos no serán considerados como un evento de determinación diferente, siempre y cuando no cambie la identidad taxonómica del ejemplar (ver. “Enmendadura”).

Las correcciones de este tipo que se hayan realizado sobre algún dato que deba ser ingresado en “Información adicional” se capturan de la siguiente forma. Las partículas no deben ser separadas con el comando de diagonales, ya que esto generaría dos partículas independientes, para indicar el orden debe ser agregada antes de la partícula original la palabra “original” y seguido de un espacio agregar la palabra “corregido” que preceda a la partícula modificada (p. ej., [...] /// original: Abundancia escaza corregido: Abundancia escasa /// [...]). Toda esta cadena de texto debe ser considerada como una partícula para efectos de la atomización en este campo.



Figura 33 Etiquetas con correcciones ortotipográficas. a) Corrección del apellido de uno de los colectores (IBUNAM:MEXU:985264); b) Corrección en el nombre de la localidad (IBUNAM:MEXU:447461); c) Corrección del género (IBUNAM:MEXU:256920).

### 7.1.3.2.8 Enmendadura

Se aplica a aquellas etiquetas que presentan tachaduras, sobre escritura, "corrector" o cualquier alteración sobre la información ya sea como intención de corrección (puede ser a mano) o eliminación.

Si presenta tachadura, se debe capturar el dato junto con el comando de corchetes más la leyenda “tachado” (p. ej., [[[F. Miranda tachado]]) en el campo correspondiente (Figura 34).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

En el caso de los eventos de determinación, la existencia de una enmendadura ocasiona que sean dos eventos diferentes si cambia la identidad taxonómica del ejemplar (p. ej., si el género, epíteto específico, epíteto infraespecífico o autoridad nuevo(s) sea diferente del evento de determinación anterior).

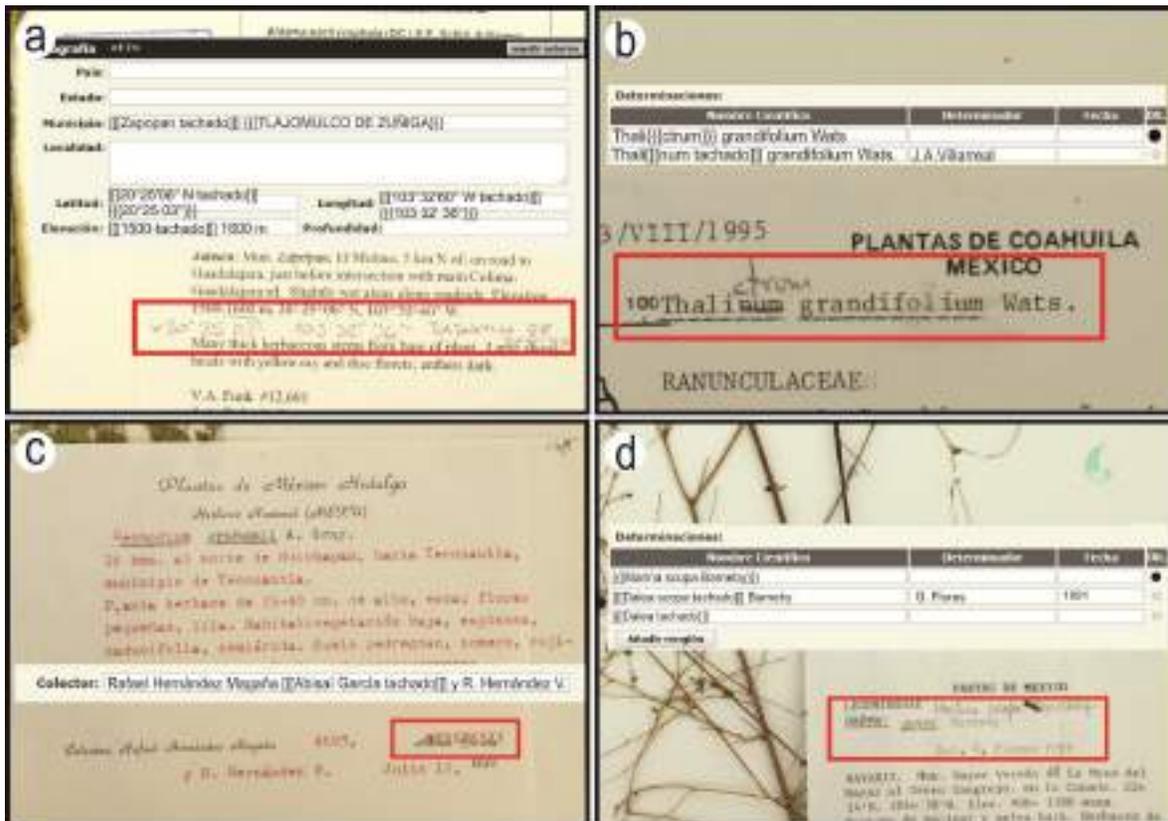


Figura 34. Etiquetas con enmendaduras. a) Tachado, se añadieron datos escritos a mano con intención de corrección (IBUNAM:MEXU:1235751); b) La corrección cambia la identidad taxonómica del ejemplar, por tanto se considera otro evento de determinación (IBUNAM:MEXU:667289); c) Tachado con intención de eliminación (IBUNAM:MEXU:285352). d) Tachado con dato de sustitución (IBUNAM:MEXU:552480).

### 7.1.3.2.9 Dañada

Aplica para etiquetas que presentan daños físicos (p. ej., rota, manchada, recortada) que impiden la lectura y captura completa de los datos (Figura 35). Es preciso realizar captura verbatim de los datos que sean legibles y sustituir lo que no por la leyenda “etiqueta dañada”, junto con el comando de corchetes (p. ej., [[[24-0-199 etiqueta dañada]]]). Esta leyenda puede ser capturada en cualquiera de los campos de la Interfaz, dependiendo de



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

la información afectada por el daño mecánico. Además, se debe informar a través de la Interfaz de Reportes de problemas en ejemplares.



Figura 35. Etiquetas dañadas. Estas etiquetas han sufrido un daño mecánico lo cual impide la visualización completa de los datos. a) Etiqueta rota que afecta a la fecha de colecta (IBUNAM:MEXU:735023); b) Etiqueta con manchas sobre varios datos (IBUNAM:MEXU:877460); c) Etiqueta recortada. El nombres del segundo colector está incompleto (IBUNAM:MEXU:1204407).

### 7.2 Reglas específicas (campos)

A continuación se presenta una serie de reglas que son específicas para cada tipo de dato, según los campos disponibles en la Interfaz de captura que aparecen en la Figura 1. Todas las reglas generales son acumulativas junto con las reglas específicas de los campos. Se muestran también una serie de imágenes asociadas a las reglas con ejemplos de captura de datos en algunos casos.

#### 7.2.1 Familia

Este campo corresponde al nombre de la familia asignada en el último evento de determinación. El nombre de la familia debe capturarse verbatim, es decir que no importa si no aparece correctamente escrito en latín (p. ej., Cactace), o si se presenta castellanizado (Cactácea), e incluso si es un nombre en desuso o incorrecto (Figura 36).



Figura 36. Ejemplos de nombres de "Familia" en etiquetas. a) Familia con denominador de campo



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

(IBUNAM:MEXU:774135); b) Sin denominador de campo y nombre castellanizado (IBUNAM:MEXU:1141132); c) Nombre de familia desactualizado (IBUNAM:MEXU:768657).

Si la familia aparece abreviada no debe completarse, sino que debe capturarse verbatim (Figura 37). Este es el único campo de la Interfaz en el cual puede capturarse la partícula “SD” en sustitución de ausencia de familia, en caso de que no esté referida la familia en la última determinación (Figura 38), no importa si alguna de las determinaciones anteriores tiene familia asociada.

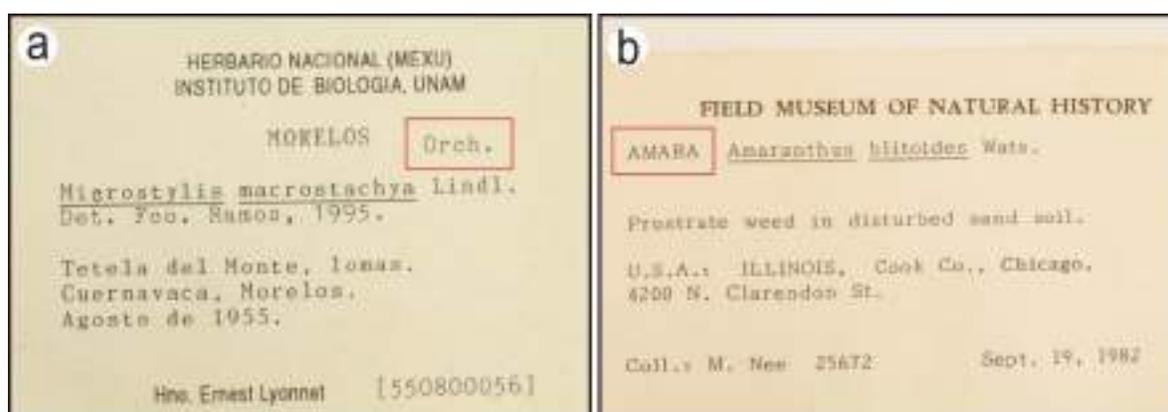


Figura 37. Etiquetas donde la familia aparece abreviada. a) Orchidaceae (IBUNAM:MEXU:657195); b) Amaranthaceae (IBUNAM:MEXU:731870).

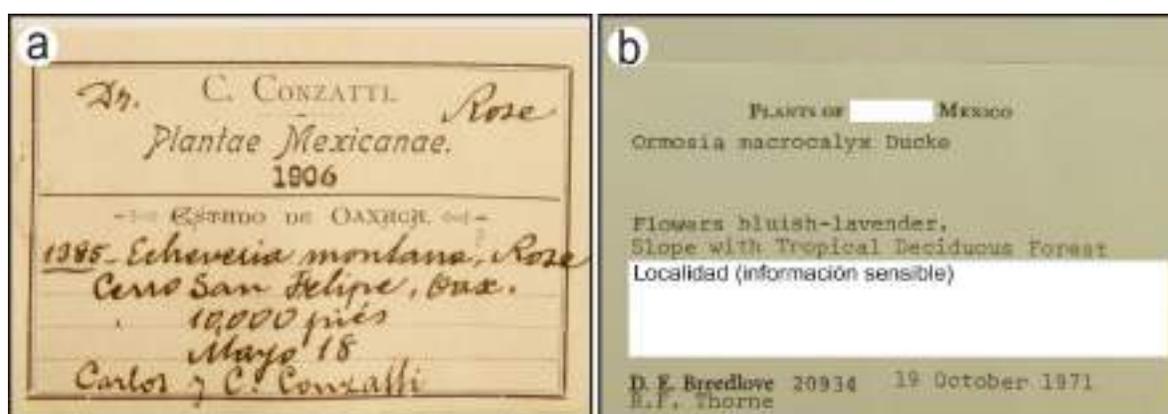


Figura 38. Etiquetas sin valor de familia. a) Pertenece a la familia Crassulaceae, pero no se indica (IBUNAM:MEXU:528754); b) Pertenece a la familia Leguminosae, pero no se especifica (IBUNAM:MEXU:199921).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.1.1 Varias familias

Cuando existen varios eventos de determinación puede existir más de un nombre de la categoría taxonómica correspondiente a familia. Cada una de las familias debe capturarse después del nombre científico al que se encuentra asociada, separándose por medio del comando de diagonales y no serán consideradas como “doble dato”, a menos que sólo exista un evento de determinación y existan dos valores de familia diferentes asociadas a dicho evento único.

Un caso especial es aquel en el que alguno de los eventos de determinación es únicamente a nivel de familia y está asociado a datos de determinador y fecha de determinación. En estos casos los datos se deben capturar en los campos correspondientes de “Nombre científico”, “Determinador” y “Fecha” respectivamente, es necesario capturar el nombre de la familia de la determinación en el campo “Nombre científico”.

Igualmente, si la única determinación del ejemplar es nivel de familia y no existe algún otro evento de determinación. El nombre de la familia se debe capturar tanto en el campo de “Familia” como de “Nombre científico”. Si existen datos sobre el determinador o la fecha de determinación asociados a este único evento, éstos deben ser capturados en sus campos correspondientes.

De modo que, el valor de familia asociado a última determinación es el único que se debe capturar en el campo de “Familia”, es decir el más reciente. Si la última determinación no tiene familia se debe de usar “SD” (Figura 39).



Figura 39. Ejemplos de determinación con varias familias asociadas. a) Existen tres eventos de determinación. El primer evento tiene una familia distinta al segundo evento, mientras que el tercer evento carece de familia (IBUNAM:MEXU:722846); b) Los tres eventos de determinación tienen asociada una familia diferente (IBUNAM:MEXU:260025).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.2 Nombre científico

El nombre científico es uno de los valores más importantes del ejemplar, pues es el que le otorga una identidad. Por lo que es imperativo que este dato sea capturado de manera verbatim. No importa si el nombre no está escrito correctamente, los autores no están correctamente abreviados, o no es el nombre aceptado.

De manera general, un nombre científico está compuesto por un género, epíteto específico y su autoridad. Sin embargo, en algunas etiquetas los nombres científicos pueden venir acompañados de otros elementos tales como el año y revista de publicación, incluso otra serie de términos propios de la colección (p. ej., que refieren al estatus de la determinación, o a la certeza de la misma). Por lo que todos los elementos asociados directamente al nombre científico deben ser capturados junto con el nombre científico en su campo.

También puede darse el caso de que el nombre científico carezca de algún elemento de los anteriormente mencionados, por lo que su ausencia debe ser respetada y no añadirse a fin de que el nombre científico sea “correcto”.

La determinación puede estar a cualquier nivel diferente a especie (p. ej., familia, subgénero, género, infraespecie, etc.). En estos casos se deben incluir en la captura todos los elementos contenidos en el nombre científico, así como los términos del rango taxonómico, que aparezcan en la etiqueta. (p. ej., subg., var., fo., subsp., cv., x, raza, etc.) y otros términos asociados (p. ej., comb. nov., aff., sensu lato, cf., complex, etc.) (Figura 40).

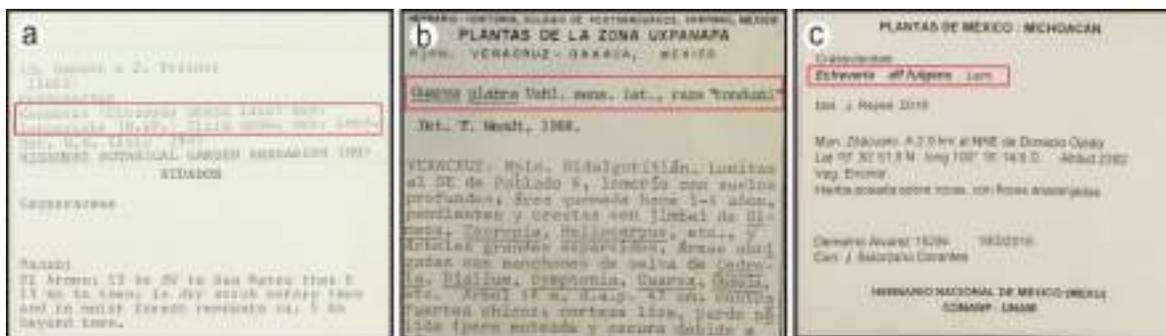


Figura 40. Ejemplos de términos asociados en determinaciones. a) La especie se indica en sentido amplio, mientras que la combinación de la intraespecie es nueva y no ha sido publicada. (IBUNAM:MEXU:482559); b) La especie se indica en sentido amplio y perteneciente a una “raza” (IBUNAM:MEXU:901129); c) La determinación del ejemplar está indicada como afín. (IBUNAM:MEXU:1449342).

De modo que cualquiera que sea el nivel de determinación debe ser capturado en el campo de “Nombre científico” como una fila diferente. Así mismo, si existen datos de determinador y fecha de determinación asociados, éstos deben ser capturados en los campos de “Determinador” y “Fecha” respectivamente (Figura 41).



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 41. Ejemplos de determinaciones en etiquetas de ejemplares. a) La determinación del recuadro rojo es sólo a nivel de Familia (IBUNAM:MEXU:722846); b) La última determinación (recuadro rojo) sólo indica la variedad del ejemplar, junto con una nota (IBUNAM:MEXU:280856); c) La determinación indica que el ejemplar debe ser confrontado para verificar la determinación (IBUNAM:MEXU:837996); d) La determinación indica que se trata de una nueva combinación de nombre científico (IBUNAM:MEXU:671394); e) Se indica que el ejemplar es afín a una especie en particular (IBUNAM:MEXU:149334); f) Se indica que el ejemplar pertenece a un “cultivar” y que pertenece a un complejo taxonómico (IBUNAM:MEXU:162352).

Puede ocurrir, en casos muy raros, que no exista ningún nivel de determinación, sólo en estos casos se debe indicar los campos de Familia y Nombre científico como “SD”, para confirmar la ausencia de cualquier categoría taxonómica. Pero son los únicos campos donde se puede indicar.

### 7.2.2.1 Nombre científico más notas taxonómicas

Cuando existan notas asociadas al nombre científico (Figura 42), la captura debe ser verbatim en el campo “Nombre científico” en el orden en que se presenten los datos (p. ej., nombre de especie, autor, publicación, variedad, autor de variedad, publicación de variedad y notas adicionales).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 42. Ejemplos de nombres científicos con notas taxonómicas. a) Se mencionan características del ejemplar (IBUNAM:MEXU:106689); b) Se incluyen publicaciones del nombre de la especie (IBUNAM:MEXU:33553); c) Se incluyen sinonimias y sus publicaciones (IBUNAM:MEXU:393951); d) Se indica el estado reproductivo del ejemplar, así como dos posibles nombres científicos (IBUNAM:MEXU:101832).

### 7.2.2.2 Diferentes eventos de determinación taxonómica

El nombre científico debe ser capturado verbatim de acuerdo a los diferentes eventos de determinación de los ejemplares, respetando el orden cronológico (comenzando por la primer determinación) siguiendo las reglas establecidas en la sección de "Nombre científico", y cualquier otra que aplique según el caso (p. ej., tachaduras, confirmaciones, etc.). La regla de "doble dato" no aplica para los eventos de determinación, aunque sean iguales. Así mismo, la presencia de enmendaduras o añadiduras genera diferentes eventos de determinación, ya que esto cambia la identidad taxonómica del ejemplar, pues una enmendadura se considera un cambio mayor que una corrección ortotipográfica, y en estos casos no se pueden asignar los valores asociados de "Determinador" y "Fecha de determinación" al nuevo evento causado por enmendadura, ya que hay certeza de quién es responsable de esa corrección (Figura 43).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

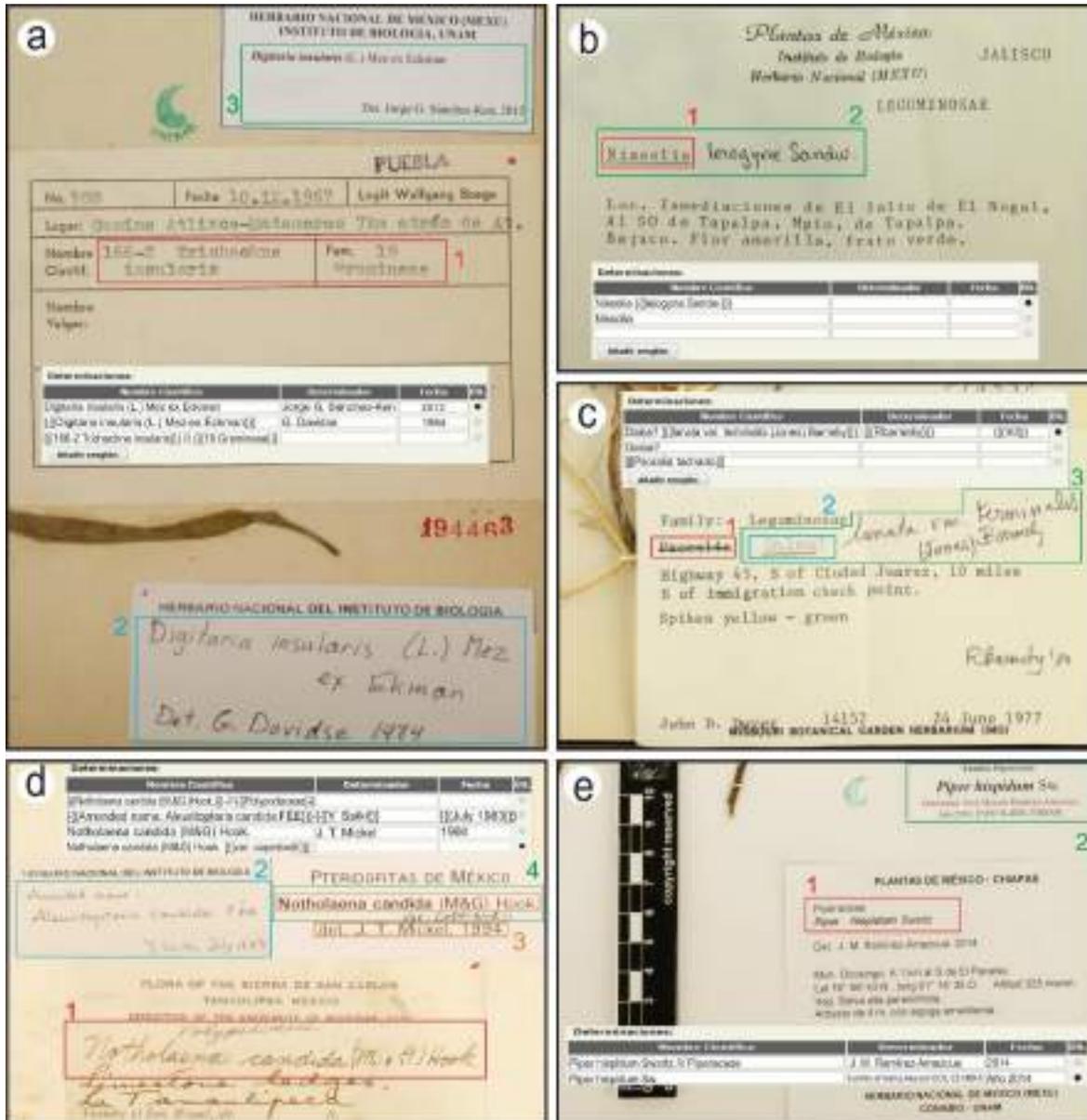


Figura 43. Ejemplos de ejemplares con varios eventos de determinación. a) Tres eventos de determinación (IBUNAM:MEXU:194463); b) Dos eventos, el último agrega el epíteto específico y el autor (IBUNAM:MEXU:743058); c) Tres eventos, el primero tachado y el segundo dudoso (IBUNAM:MEXU:274852); d) Cuatro eventos de determinación. La última etiqueta presenta dos eventos (IBUNAM:MEXU:88252); e) Son dos eventos de determinación aun cuando prácticamente sean iguales los datos (IBUNAM:MEXU:1404830).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.2.3 Confirmación de nombre científico

Las confirmaciones taxonómicas, normalmente representadas por un signo de admiración (!), expresan que el nombre científico inmediato anterior (cronológicamente), es correcto y se toma como una determinación válida, por tanto las confirmaciones son consideradas como un evento de determinación (Figura 44).

Estos eventos se deben capturar verbatim, según aparezca referida la confirmación del dato existente. Se deben capturar el o los signos de admiración en el campo correspondiente al elemento del nombre científico al que están asociados, ya sea al nombre científico, al determinador o fecha de determinación; así mismo se debe capturar el texto asociado o el denominador de campo presente, por ejemplo: “identificación correcta”, “conf.”, en opciones de verificación (Figura 44 inciso d) (ver. “Opciones múltiples, de llenado o marcado”), etc.

En los casos donde solamente se encuentran el o los signos de admiración, se debe capturar el nombre científico y la autoridad de la determinación con fecha inmediatamente anterior a la confirmación, tal y como aparece en dicha determinación; esto incluye todos los elementos asociados al nombre científico, no así el nombre del determinador y fechas anteriores (Figura 44 inciso a y b).



Figura 44. Ejemplos de ejemplares con eventos de determinación indicados como confirmaciones. a) La última determinación está indicada como una confirmación (IBUNAM:MEXU:29729); b) La última determinación es una confirmación y no se encuentra en una etiqueta independiente (IBUNAM:MEXU:372748); c) El último evento de determinación sólo indica las iniciales de



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

determinador (IBUNAM:MEXU:801691); d) La etiqueta de determinación indica la opción “conf.” (IBUNAM:MEXU:1265884); e) Se indica la determinación por medio de la expresión latina “Determinavit” (IBUNAM:MEXU:717626); f) Se indica confirmación y se incluyen el nombre científico y la familia (IBUNAM:MEXU:154692).

### 7.2.2.4 Números asociados a taxa

Algunas etiquetas contienen números asignados a alguna de las categorías taxonómicas (p. ej., núm. de la familia o género). Estos números corresponden a los empleados por distintos tipos de clasificaciones taxonómicas y no deben confundirse con números de colecta, ya que generalmente corresponden a un evento distinto al de colecta y en algunos casos se indica con un denominador de campo (ver “Número de colecta”). La captura de estos números debe ser verbatim en el campo correspondiente según el dato al que estén asociados (Figura 45). Deben además, distinguirse de los números sin asociación, ya que generalmente están escritos en un evento diferente y fuera de la etiqueta (ver. “Números sin asociación”).



Figura 45. Etiquetas con un número asignado a taxa. a) Número asignado al denominador familia, sin el nombre de ésta (IBUNAM:MEXU:392662); b) Dos números asignados a una familia, uno de ellos se encuentra tachado y un valor numérico asignado a nombre científico. Estos números pueden distinguirse del número de colecta por los denominadores de campo (IBUNAM:MEXU:96218); c) Número y nombre de la familia asignados al campo correspondiente (IBUNAM:MEXU:1470498).

### 7.2.3 Determinadores y colectores

Al capturar los nombres de los determinadores y colectores (personas, equipos o instituciones), se debe capturar de manera adjunta los nombres, acrónimos, abreviaturas, iniciales o siglas de la institución de adscripción de dichas personas, los nombramientos laborales o académicos y cualquier partícula asociada (Figura 46). Algunos denominadores de campo poco comunes para estos nombres incluyen: legitor (leg. o legit.) para colectores; y Fide o Vidit para determinadores. En el caso de los determinadores, la captura sigue un orden cronológico determinado por las fechas de determinación (ver “Nombre científico”).



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 46. Determinadores y colectores. a) Denominador de campo (Legit) y Nombre de colector (IBUNAM:MEXU:194463); b) El nombre del colector corresponde a un proyecto (IBUNAM:MEXU:767032); c) Se incluye nombramiento del colector (IBUNAM:MEXU:956970); d) Incluye siglas del herbario de adscripción y proyecto por el cual fue revisado el ejemplar (IBUNAM:MEXU:621737); e) Incluye doble dato de colector (un equipo y el colector principal) (IBUNAM:MEXU:52387); f) Doble dato de colector, se incluye para quien fue colectado y por quien (IBUNAM:MEXU:21323).

### 7.2.3.1 Firma

En algunas etiquetas los determinadores o colectores pueden escribir su firma en lugar de su nombre para indicar que son ellos quienes determinan o colectan un ejemplar. (Figura 47). Se debe capturar la leyenda “firma” en el campo correspondiente, así mismo se debe



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

usar comando de llaves si es necesario (p. ej., {{{{firma}}}}). No debe confundirse con nombres ilegibles.



Figura 47. Etiquetas de revisión taxonómica con la firma del determinador. Estas firmas pueden ser investigadas en procesos posteriores a la captura, pero esta regla permite su rastreo dentro de la base de datos. a) La firma corresponde Fraga (IBUNAM:MEXU:131816); b) El determinador corresponde a Terence Dale Pennington (IBUNAM:MEXU:573496).

### 7.2.4 Fechas de determinación y colecta

En la captura de las fechas, ya sean de determinación o colecta, se debe respetar el formato que presenten las etiquetas. Por lo que no deben completarse los valores abreviados, ni cambiar los meses escritos con letra a números, ni cambiar el orden de los elementos de la fecha o sus separadores (p. ej., cambiar “sep 3 -18” a “3/9/2018”) y mucho menos agregar connotaciones para indicar la ausencia de algunos de los elementos de fecha, por ejemplo esta notación: 00/00/2000, trata e indicar que sólo hay año (Figura 48).

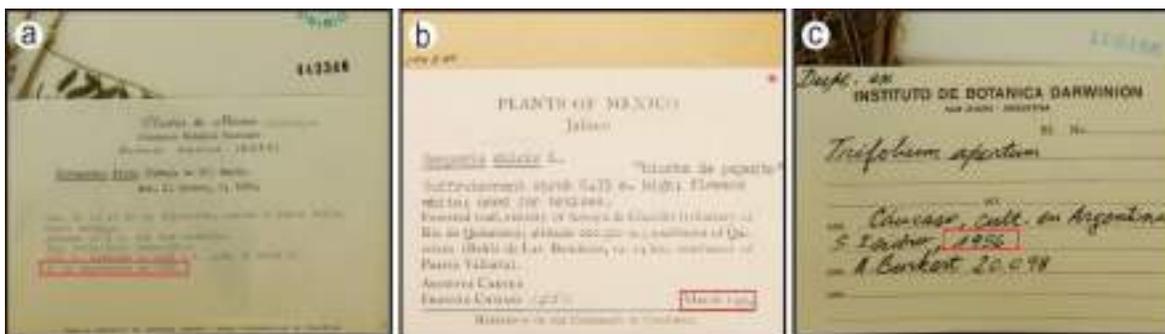


Figura 48. Ejemplos de formatos de fechas. a) La fecha incluye el artículo “de” dos veces, el mes en formato completo y el año (IBUNAM:MEXU:443346); b) La fecha está en inglés y no indica el



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

día de colecta (IBUNAM:MEXU:106340); c) Sólo se indica el año de colecta (IBUNAM:MEXU:169486).

### 7.2.5 Número de colecta

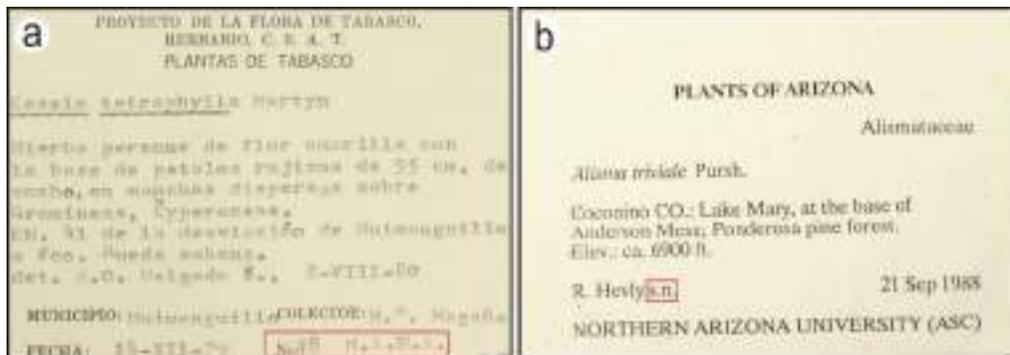
Se debe capturar el número o dato indicado en el denominador de campo específico de la etiqueta (cuando esté presente) o el número asociado a nombres, ya sea de personas, equipos, proyectos, etc., que hagan alusión al responsable de la colecta del ejemplar (Figura 49). Si el número de colecta incluye iniciales del colector, siglas de algún proyecto o cualquier otro indicador especial, éstos deben ser capturados junto con el número de colecta. Lo mismo ocurre si el denominador de campo indica el número de colecta como “s.n.” (sin número, o cualquiera de sus variantes). Sin embargo, no se debe usar esta notación para indicar la ausencia del número de colecta.

En el caso de equipos de colecta, en los que uno o más de éstos tengan un número de colecta (igual o diferente), se deben capturar todos los nombres del equipo de colecta con sus respectivos números asociados en el campo de “Colector” y en el campo de “Número de colecta” se debe capturar solamente el valor del número de colecta del primer colector del equipo.

En algunas etiquetas antiguas, el número de colecta aparece a la izquierda del nombre científico, incluso con un sello. Sin embargo, para poder capturar dicho número en el campo correspondiente, deben cumplirse las siguientes condiciones, de lo contrario se debe capturar en el campo de “Información adicional” (ver. “Números sin asociación”):

- 1.- El número debe estar escrito dentro de la etiqueta principal de colecta.
- 2.- No debe existir algún otro número con o sin denominador de campo que indique que es otro el número de colecta.
- 3.- Debe estar escrito en el mismo evento, y no dar la apariencia de haber sido añadido después.

Dos ejemplos de estos casos se muestran en los incisos e y f de la Figura 49.



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

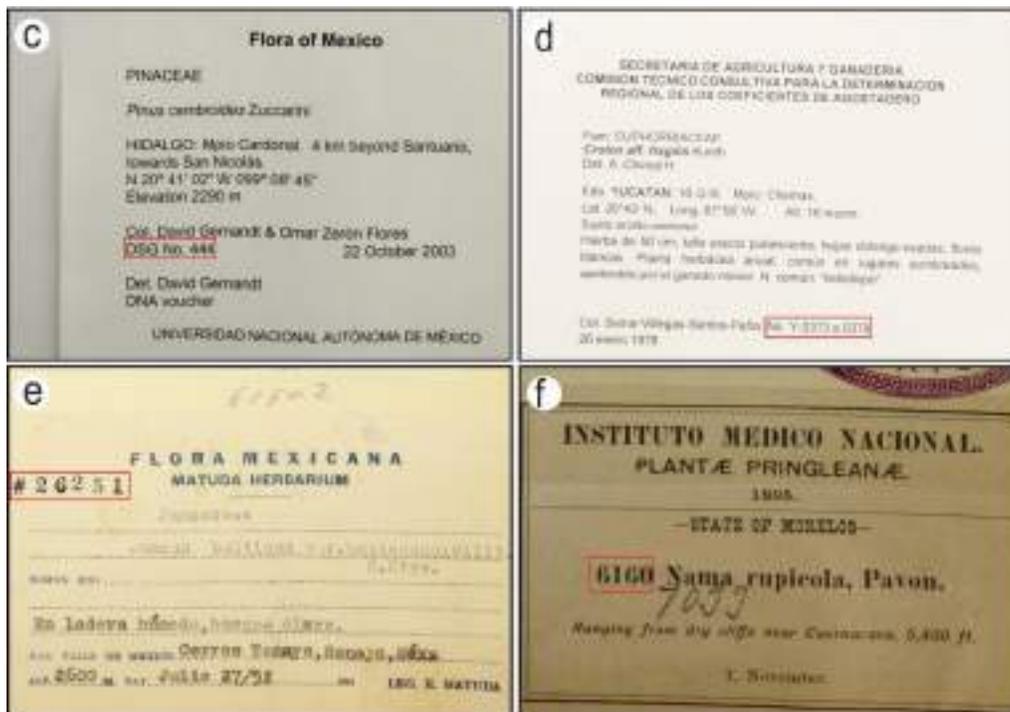


Figura 49. Números de colecta en etiquetas. a) El número incluye las iniciales del colector (IBUNAM:MEXU:292122); b) El número se indica como "s.n." (IBUNAM:MEXU:26405); c) El número incluye las iniciales del colector (IBUNAM:MEXU:1244887); d) Se indica una serie de números de colecta (IBUNAM:MEXU:1110032); e) El símbolo de "gato" indica el número de colecta (IBUNAM:MEXU:81802); f) El número de colecta es el que aparece junto al nombre científico, el número bajo el nombre del género no es número de colecta, además está escrito en un evento posterior al de la colecta (IBUNAM:MEXU:25515).

### 7.2.6 País, estado, municipio y localidad

Estos datos geográficos se deben capturar de forma total y verbatim. Debe incluirse el nombre de la división administrativa de estado y municipio, cuando son distintas a la de la Interfaz de captura y respetarse el formato que éstas presenten, según sea el caso (p. ej., Provincia, Departamento, Prov., Depto., Cantón, Distrito, etc.). Si existe "doble dato", ambos deben capturarse en el campo correspondiente, separándolos por medio del comando de diagonales (Figura 50).

Si existen más de dos divisiones administrativas por debajo de estado, la segunda debe capturarse en el campo de "Localidad" junto con la localidad, pero debe seguirse el orden de las oraciones establecido en la etiqueta, para poder asignar correctamente la información.



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

En el caso particular de ejemplares de México y del estado de Oaxaca, los datos geográficos con frecuencia incluyen niveles adicionales de división administrativa no oficial (p. ej., Distrito y Región). Estos nombres y sus denominadores deben ser capturados en el campo de “Localidad”, en el orden en que aparecen en la etiqueta y respecto de la localidad.



Figura 50. Información geográfica en etiquetas. Se indica la división administrativa en recuadros rojos y en azul las localidades. a) Existe doble dato para el campo de “Localidad” (IBUNAM:MEXU:217628); b) Se indican país, estado y localidad (IBUNAM:MEXU:1055521); c) Existen tres divisiones administrativas, las primeras dos se deben asociar al campo de “Estado” y “Municipio”, respectivamente (IBUNAM:MEXU:607084); d) Existen tres divisiones administrativas, el nombre bajo el denominador de campo de “Municipio” debe ser capturado en el campo correspondiente, mientras que los otros dos nombres en el campo de “Estado”, siguiendo la regla de “Asignación de información asociada” y puede agregarse la leyenda “confuso” con el comando de corchetes (IBUNAM:MEXU:1397466); e) El nombre “DISTRITO FEDERAL” no tiene denominador de campo, pero puede ir en el campo de “Estado” (IBUNAM:MEXU:721542); f) La división administrativa por debajo de país, no es clara si es que se trata de dos Prefecturas o de otro nivel distinto, puede capturarse la leyenda “confuso”, junto con el comando de corchetes en el campo de “Estado”(IBUNAM:MEXU:1181980).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.7 Coordenadas

Se deben capturar verbatim, tomando en cuenta no omitir las partículas asociadas a ellas (p. ej., “ca”, “about”, “aprox.”, etc.). Al realizar la asignación no se deben descontextualizar los datos o cambiar su sentido (ver “Asignación de información asociada” en la sección de “Asignación y atomización de la información”). Se debe usar el comando de diagonales en caso de existir “doble dato”, por ejemplo cuando existen dos pares de coordenadas o si existen coordenadas en dos formatos diferentes (Figura 51). Para indicar los minutos (') y segundos (") se deben usar comillas simples. No debe agregarse el sentido de la orientación de las coordenadas si no aparecen indicadas, por ejemplo dirección Norte (N) y W (Oeste) respectivamente.

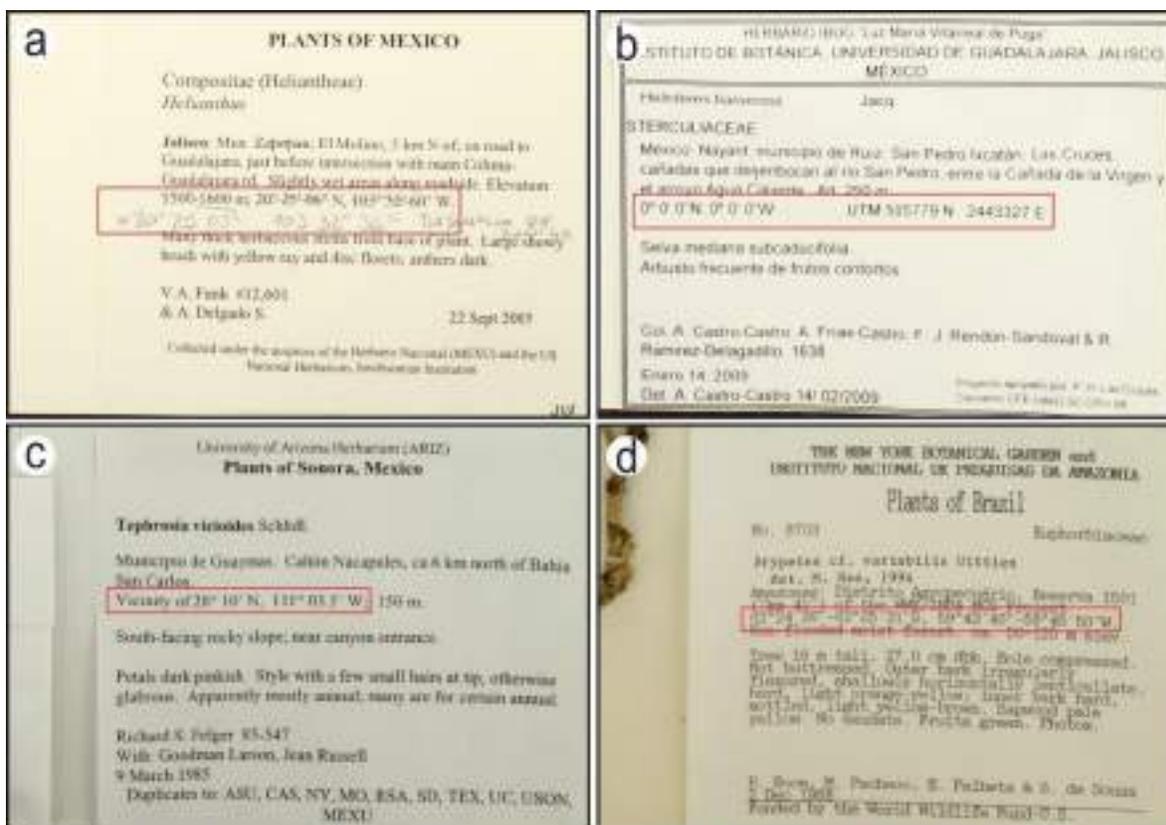


Figura 51. Ejemplos de etiqueta con coordenadas. a) En formato de grados, minutos y segundos con tachadura y añadidura escrita a mano (IBUNAM:MEXU:1235751); b) Doble dato y formato. Se indican en grados, minutos y segundos y en UTM (IBUNAM:MEXU:1371860); c) Con la partícula “Vicinity of” (IBUNAM:MEXU:1350811); d) Rango de coordenadas en formato de grados, minutos y segundos (IBUNAM:MEXU:660217).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.8 Elevación

La elevación debe ser capturada verbatim, incluso si el valor designado como elevación no es una cantidad (p. ej., a nivel del mar). No deben omitirse partículas asociadas a la elevación (p. ej., “ca.”, “about.”, “aprox.”, etc.). Si la elevación se presenta en forma de intervalos se debe mantener su presentación y no descontextualizar o atomizar los datos. Sin embargo, si se presenta “doble dato” deben separarse por medio del comando de diagonales (Figura 52).



Figura 52. Ejemplos de datos de elevación en etiquetas. a) Se muestra un intervalo de elevación con una partícula de incertidumbre (IBUNAM:MEXU:159130); b) Se muestra un intervalo con aproximación (IBUNAM:MEXU:1277982); c) La elevación se muestra como una descripción (IBUNAM:MEXU:429594).

### 7.2.9 Forma de vida

Los datos a capturar en este campo serán los presentes en el denominador de campo de forma de vida de las etiquetas y en caso de no existir, se debe limitar al vocabulario controlado creado para este campo. A continuación se muestra una lista con las formas de vida más comunes y algunas de las variantes encontradas hasta el momento (Tabla 1). La lista está sujeta a actualización conforme se encuentren más variantes.

Tabla 1. Vocabulario controlado para el campo de “Forma de vida”.

Forma de vida	de	Variantes y similares
Adulto		
Anual		annuelle
Árbol		arbolillo, arbolito, arbor, arbórea, arbóreo, arborescent, arborescente, arbre, arbrè, arvore, arvoreta, baum, shrubby tree, tree, treelet, tree-like
Árbol o Arbusto		arb.
Arbustiva		arbustosa, bushy, fruticosa, shrub-like, shrubby

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Forma de vida	Variantes y similares
Arbusto	arbrisseau, arbustito, bush, shrub, shrublet, strauch, undershrub
Bejuco	abejucada, bejucoide, bejucosa
Bianual	biennial
Caducifolia	caduca, decídua, deciduous
Caméfito	camephyta
Criptófita	
Enredadera	creeper, enredadero, schlingpflanze, trepadeira, twiner, twining, vine
Fanerófita	
Geófita	
Helecho arborescente	helecho arbóreo, tree fern
Hemicriptófita	
Hemigeófita	hemigeófito
Hierba	erva, forb, herb, herbácea, herbaceous, herbe, hierbita, kraut, yerba
Juvenil	joven, juvenile, plántula
Liana	liane, lianescente, volúvel
Pasto	grass
Perenne	perennial, vivace
Perennifolia	evergreen
Semiarbusto	halbstrauch, semiarbustiva
Subherbácea	subherbácea
Sufrútice	shrubby herb, sous-arbrisseau, subarbustiva, subarbusto, subfruticosa, sub shrub, suffretex, suffrutescent, suffrutescente, suffruticoise, sufruticosa
Terófita	
Viejo	old
Zacate	

### 7.2.10 Hábitat

Se trata de la categoría o descripción del sitio donde se colectó el ejemplar, que puede incluir el tipo o descripción del suelo.

Generalmente se encuentra descrito como tipo de vegetación establecido, por ejemplo: Bosque de pino-encino, Matorral xerófilo, Selva baja caducifolia, etc. También



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

Sugerencia de cita: DGRU.2020. Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2 (Versión 1). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

puede encontrarse como descripciones más extensas que incluyen elementos adicionales, por ejemplo: Ladera en bosque de pino; Selva mediana subperennifolia, secundaria, suelo café arenoso; Ecotonía de Selva baja Caducifolia y Bosque de pino-Encino; Cerca de la cima bosque de pino-encino en laderas S mesófilo con *Cedrela*, *Ulmus mexicana*, *Oecopetalum*, *Quercus*, *Podocarpus*, *Calophyllum*; etc.

No debe usarse el comando de diagonales para separar partículas en este campo, a menos que se trate de un “doble dato”.

La información contenida en el denominador de campo hábitat debe respetarse, a pesar de que contenga información adicional o que no sea propia de la naturaleza del campo. Lo mismo ocurre si la información está escrita en otro idioma y la extracción de datos puede resultar en la alteración de la información, la captura debe ser verbatim en el campo “Habitat”.

Por otro lado, si ocurre el caso en que la información bajo el denominador hábitat corresponda en su totalidad a la localidad, esta información debe ser capturada en el campo de “Localidad” incluyendo el denominador original (p. ej., Hab, Ambiente, etc.).

La Figura 53 muestra algunos de los ejemplos mencionados.



Figura 53. Ejemplos de información de hábitat en etiquetas. a) Se muestra como una categoría de tipo de vegetación (IBUNAM:MEXU:1358588); b) Se muestra bajo el denominador “Ecol.” (IBUNAM:MEXU:657439); c) La descripción incluye vegetación asociada (IBUNAM:MEXU:183416); d) La información del denominador “HAB.” corresponde íntegramente a la localidad de colecta (IBUNAM:MEXU:108863); e) La oración que corresponde a hábitat no está asociada al



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

denominador correspondiente, sino que es la localidad la que se asocia con éste. Sin embargo, las oraciones pueden asignarse a sus campos correspondientes fácilmente sin descontextualizar la información (IBUNAM:MEXU:681119); f) El hábitat está escrito en latín, pero contiene información combinada de localidad y hábitat. Debe capturarse en el campo “Hábitat” a fin de evitar la descontextualización o asignación errónea de la información (IBUNAM:MEXU: 829903).

### 7.2.11 Características del ejemplar

Es información que describe al ejemplar. Puede incluir medidas (p. ej., longitud en metros = 0.014; peso en gramos = 120; de 1.5m de alto, etc.) y características cualitativas (p. ej., corteza rojiza; flores pequeñas blancas, dispuestas en racimo; hojas compuestas alternas, etc.). No deben separarse las características con el comando de diagonales, a menos que se trate de “doble dato” (Figura 54). Únicamente en el caso en que se describa el estado reproductivo (p. ej., con flor), pero sin estar asociado a un adjetivo que describa la condición, dicha información debe ser capturada en el campo de “Información adicional” como una partícula independiente (ver. “Condición reproductiva” en la sección de “Información adicional”).

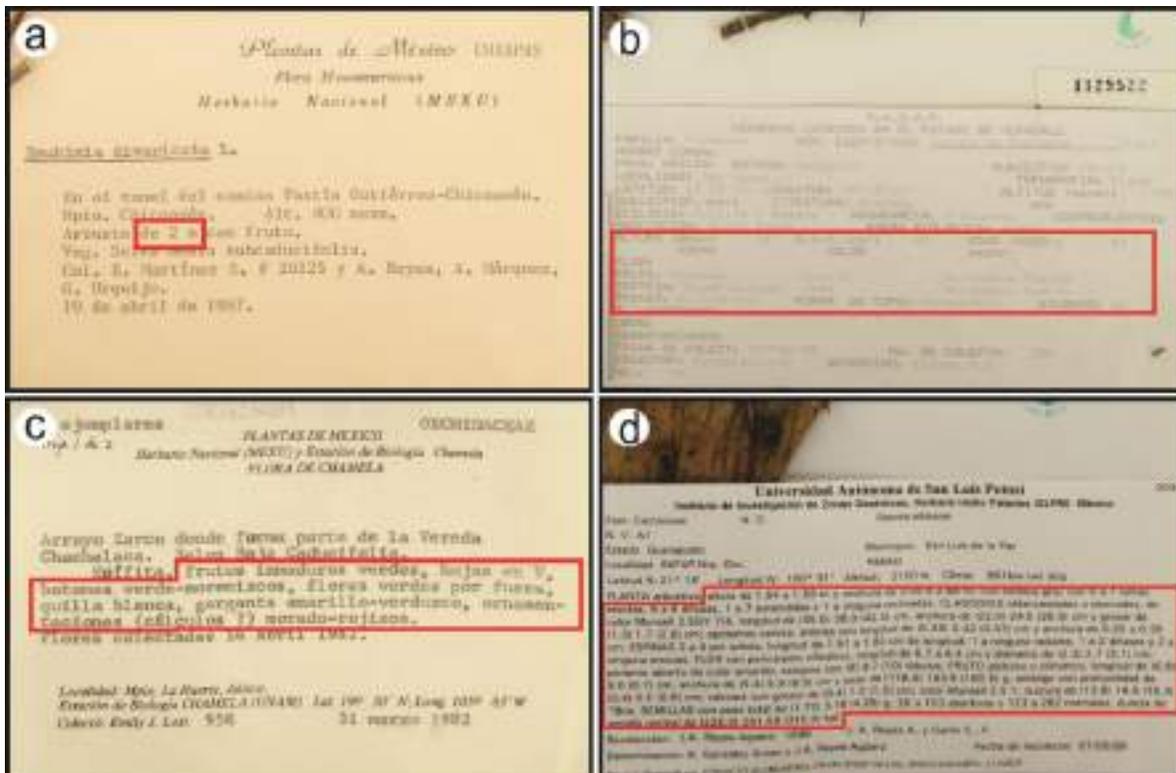


Figura 54. Etiquetas con información correspondiente a características del ejemplar. a) Se muestra la altura de la planta. “Con fruto” corresponde a “Condición reproductiva” por lo que no se captura



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

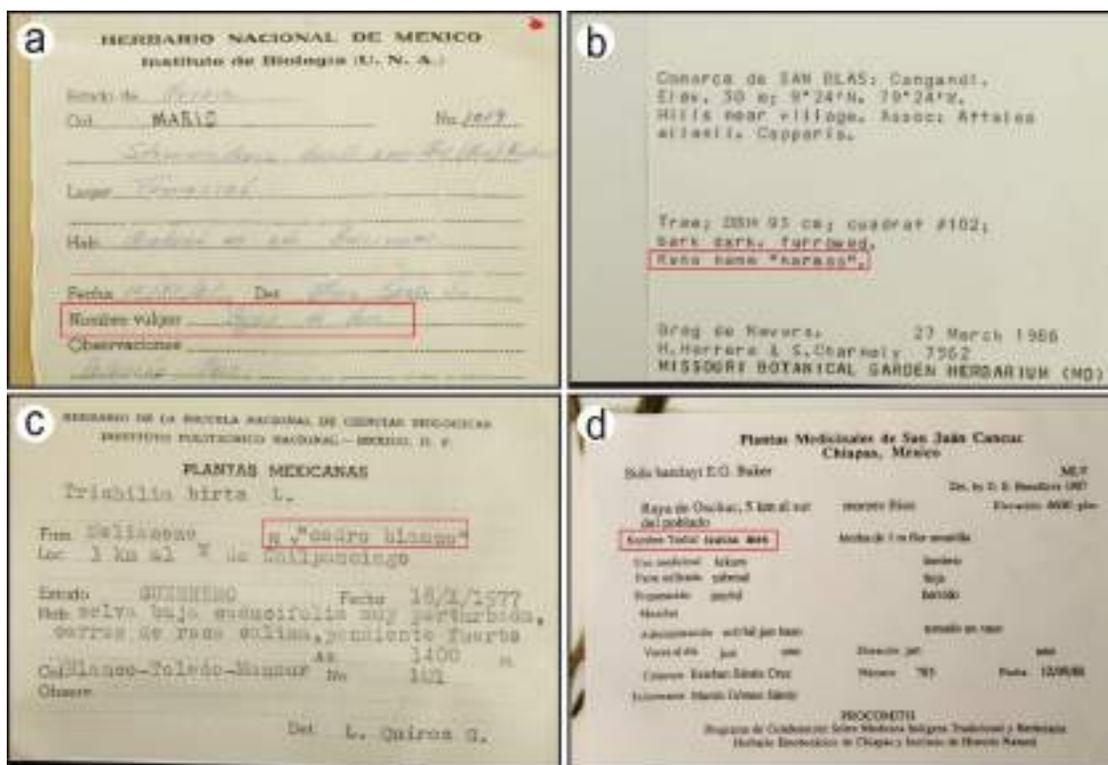
FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

en el campo de “Características del ejemplar” si no en “Información adicional” (IBUNAM:MEXU:547460); b) Casi todas las características tiene un denominador de campo propio, que igualmente deben capturarse (IBUNAM:MEXU:112952); c) Las características corresponden a una lista sin denominador de campo. “Epífita” corresponde a “Hábito” y se captura en el campo de “Información adicional” (IBUNAM:MEXU:332298); d) A pesar de ser un texto extenso debe capturarse toda la información, incluyendo los denominadores de campo dentro del texto y toda la puntuación (IBUNAM:MEXU:1214267).

### 7.2.12 Nombre común

Esta información corresponde al nombre vulgar, vernáculo, popular o local asignado a la planta. Puede presentarse en varios idiomas distintos al español o alguna lengua indígena (p. ej., tzotzil, kuna, etc.) e incluso tener varias traducciones al mismo tiempo (Figura 55). La captura debe ser verbatim, en el campo correspondiente. Debe respetarse el orden en que aparezcan los nombres e incluir el idioma o lengua al que corresponden si está indicado.

En algunos casos, puede existir dentro de la etiqueta, alguna palabra entre comillas, sin embargo, si no está asociado a un denominador de campo claro, esta información debe capturarse en el campo de “Información adicional”.



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Figura 55. Información sobre nombres comunes en etiquetas. a) Denominador de campo es “Nombre vulgar” (IBUNAM:MEXU:113117); b) Denominador de campo es “Kuna name” (IBUNAM:MEXU:706691); c) El denominador de campo está abreviado “N. v.” (IBUNAM:MEXU:842601); d) El denominador de campo indica la lengua a la que pertenece el nombre común (IBUNAM:MEXU:1145962).

### 7.2.13 Información adicional

En este campo se deben capturar todos los datos de las etiquetas que no tienen asignado un campo en la Interfaz de captura, así como los datos e información cuya naturaleza no sea clara y comprensible. Además de los datos estructurales de las etiquetas y del ejemplar en general, tales como encabezados, pies, sellos, notas accesorias, etc.

Los datos se deben separar en partículas de información coherentes, separándolas temáticamente con el comando de diagonales. La atomización de los datos debe atender a que no se pierda el sentido de la oración o se descontextualice la información, incluyendo los denominadores de campo correspondientes. Todas las partículas capturadas en este campo deben seguir las reglas de captura generales según corresponda. (Figura 56).

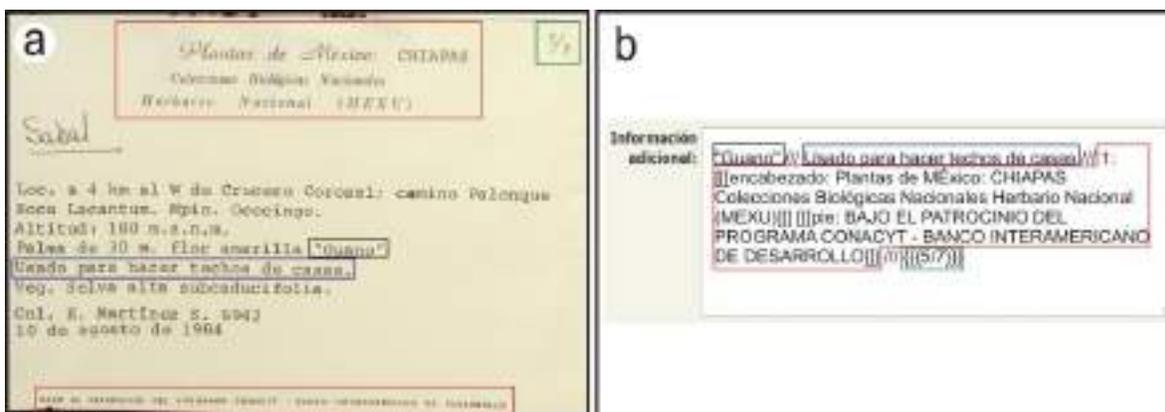


Figura 56. Ejemplo de datos que se capturan en “Información adicional” (IBUNAM:MEXU:1166200). a) Vista de los elementos de la etiqueta que corresponde al campo de “Información adicional”; b) Ejemplo de captura de dichos elementos.

Algunos de los tipos de datos que deben ser capturados en este campo incluyen:

**7.2.13.1 Notas de campo.**- son las notas tomadas en campo sobre el evento de colecta.

**7.2.13.2 Factores ambientales.**- durante el evento de colecta pueden describirse condiciones ambientales como: día soleado, con lluvia, temperatura agradable, a la sombra, etc.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

**7.2.13.3 Valores promedio anual o total.**- tales como precipitación o temperatura (p. ej., 455 mm, temperatura media anual de 28 ° C, etc.).

**7.2.13.4 Clima.**- tipos de clima ya sea en claves o descripción (p. ej., cálido húmedo, seco, templado, Aw, BS, BW, etc.).

**7.2.13.5 Condición reproductiva.**- describe la fenología del organismo o su estado reproductivo. Generalmente no incluyen adjetivos que describan la parte reproductiva. (p. ej., con flor, con fruto, sin flor, sin fruto, en floración, en fructificación, fruto inmaduro o maduro, flor en botón, flor, fruto, vaina, flor: inflorescencia, vegetativa, con soros, etc.).

**7.2.13.6 Taxa asociados.**- estos pueden incluir una lista de especies asociadas con denominador de campo separado de hábitat u oraciones concretas como visitado por mariposas, epífita de *Pinus*, etc.; es importante que si estas listas se atomizan o asignan a partir de una oración sin denominador de campo, no se descontextualice la información.

**7.2.13.7 Abundancia.**- pueden ser categorías o descripciones (p. ej., escasa, un solo individuo, abundante, ejemplar visto, abundante en las laderas y escaso en la cima, etc.).

**7.2.13.8 Usos.**- como medicinal, ornamental, alimenticio, frutal, etc.; así como modo de preparación.

**7.2.13.9 Medios de establecimiento.**- generalmente expresados por categorías (p. ej., nativa, exótica, invasiva, cultivada, silvestre, introducida, etc.).

**7.2.13.10 Hábito y comportamiento.**- de forma similar a forma de vida, estos valores pueden estar bajo denominador de campo correspondiente, de no ser el caso, en la Tabla 2 se muestra una lista de vocabulario controlado para este tipo de dato.

Tabla 2. Vocabulario controlado para hábito y comportamiento.

Hábito y comportamiento	Variantes y similares
Apoyada	apoyante
Arvense	
Ascendente	ascending
Colgante	
Decumbente	decumbent, dressée
Epífita	epífito, epiphyte, epiphytic
Erecta/Erguida	erect, erguido



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Hábito y comportamiento	Variantes y similares
Escandente	clambering, scandent
Estrangulador	strangler
Hemiepífita	hemiepiphyte
Hemiparásita/Semiparásita	Hemiparásito
Hidrófita/Acuática	acuática arraigada de hojas flotantes, acuática emergente, acuática flotante, emergente, enraizada, enraizado emergente, flotante, hidrófita enraizada con hojas flotantes, hidrófita enraizada emergente, hidrófita libremente flotadora, subacuática, sumergida
Inclinada	leaning herb
Litófita	saxícola
Maleza	weed, weedy
Parásita	
Procumbente/Prostrada	postrado, prostrate, procumbent, prostrée, recargada, reclining, tendido
Rastrera	creeping, rampante, repent
Rizomatosa	rhizomatous
Ruderal	
Rupícola	epipétrico
Saprófita	
Semiacuática	
Semierecta/Semierguida	semi - erecta
Semipostrada	semi - postrada, semipostrado
Semirastrera	
Semitrepadora	semitrepador
Subpostrada	subpostrate
Terrestre	terrestrial
Trepadora	climber, climbing, scrambling, trepador, vining

Otros elementos que deben ser capturados en el campo de “Información adicional”, incluyen: elementos estructurales de las etiquetas principales y secundarias, como los encabezados y pies; sellos propios de las etiquetas o del ejemplar mismo (impresos en el montaje); etiquetas con texto asociado; etiquetas de nota accesoria;



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

Sugerencia de cita: DGRU.2020. Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2 (Versión 1). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

anotaciones curatoriales y números sin asociación. A continuación se describe cada uno de estos elementos y la forma en que deben ser capturados.

### 7.2.13.11 Encabezados y pies de etiquetas principales y secundarias

Dentro de la diversidad de etiquetas se encuentran aquellas que presentan encabezados y pies, los cuales contienen información que también debe ser capturada (Figura 57).



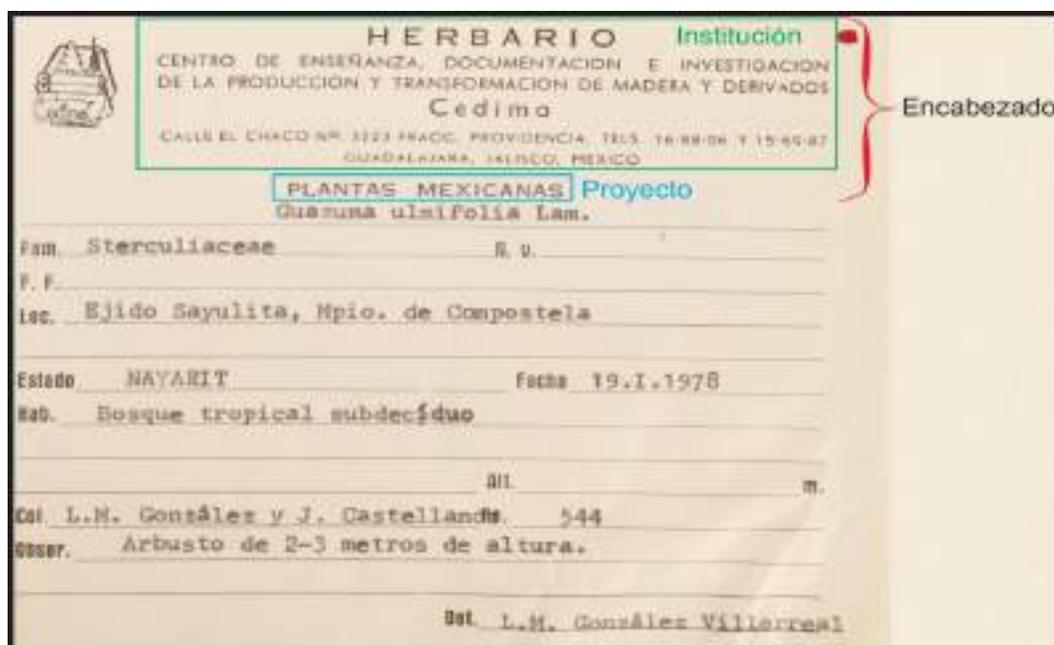
## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 57. Ejemplos de encabezados y pies de etiquetas de colecta. a, b, c y d) Encabezados de etiquetas (IBUNAM:MEXU:1074297, IBUNAM:MEXU:100225, IBUNAM:MEXU:100543 y IBUNAM:MEXU:1008614, respectivamente); e, f, g y h) Pies de etiquetas (IBUNAM:MEXU:1014539, IBUNAM:MEXU:584548, IBUNAM:MEXU:1009349, IBUNAM:MEXU:1011637, respectivamente).

Algunos de los tipos de información que pueden incluir los encabezados y pies de etiquetas son: nombres de instituciones que llevaron a cabo los proyectos de colecta, que los respaldan o fungen como repositorios de los ejemplares; así como los mismos nombres de los proyectos y de sus financiadores (Figura 58). Si bien este tipo de información no tiene correspondencia con los campos de la Interfaz de captura, en los análisis posteriores a la captura han resultado ser útiles para la identificación de bloques de ejemplares con información en común, por ello es importante que sean capturados.



**HERBARIO** Institución  
CENTRO DE ENSEÑANZA, DOCUMENTACIÓN E INVESTIGACIÓN  
DE LA PRODUCCIÓN Y TRANSFORMACIÓN DE MADERA Y DERIVADOS  
Cedima  
CALLE EL CHACO Nº. 3223 FRACC. PROVIDENCIA, TELS. 16 88 06 Y 15 65 87  
GUADALAJARA, JALISCO, MÉXICO

PLANTAS MEXICANAS Proyecto  
*Quaruna ulmifolia* Lam.

Fam. Sterculiaceae R. V.  
F. F.  
Loc. Ejido Sayulita, Mpio. de Compostela  
Estado NAYARIT Fecha 19.I.1978  
Hdb. Bosque tropical subdeciduo  
Col. L.M. González y J. Castellanos 544  
Göser. Arbusto de 2-3 metros de altura.  
Det. L.M. González Villarreal

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Figura 58. Ejemplo del tipo información presenten en un encabezado. En este caso una “Institución” y un “Proyecto” (IBUNAM:MEXU:262408).

Dado que la información que contienen así como su presentación es tan variada, dentro de las etiquetas, se dificulta definir y delimitar ambos conceptos. De modo que a falta de un denominador de campo claro para delimitar estos elementos, la captura de los datos que contengan se debe realizar de acuerdo a las siguientes instrucciones.

Se deben capturar los encabezados y pies de todas las etiquetas en orden cronológico y de tipo de etiqueta. Se deben asignar números consecutivos a cada una, comenzando por la etiqueta principal del de evento de colecta con el número uno (1), le siguen las etiquetas de determinación (en orden cronológico respecto a la fecha del evento) y las etiquetas “Adicionales al evento” (en orden arriba a abajo y de izquierda a derecha), en caso de que éstas tengan encabezados o pies; se continúa así hasta enumerar y capturar todas las etiquetas. Respecto a las etiquetas de “Nota accesoria” o de “Texto asociado”, a pesar de que contengan encabezados o pies, éstos no deben capturarse independientemente, ya que tienen otra instrucción para ser capturadas (ver. “Diversidad de etiquetas”).

La notación de captura es la siguiente: escribir el número de etiqueta (1, 2, 3, etc.), seguido de dos puntos (:); después usar el comando de corchetes para escribir la palabra “encabezado”, seguido de dos puntos y el texto del encabezado; si existe también un pie, dejar un espacio sencillo después del triplete de corchetes de cierre y usar nuevamente el comando de corchetes para escribir la palabra “pie”, seguido de dos puntos y el texto del pie de etiqueta. Toda esta cadena de texto es considerada como una partícula única, por lo que debe ser separada de otras partículas con el comando de diagonales, de ser necesario.

Si una etiqueta sólo tiene encabezado o pie solo se indica el que presente, sin precisar la ausencia del otro. Igualmente si el encabezado o pie están total o parcialmente escrito a mano debe indicarse por medio del comando de llaves (p. ej., 1: [[[encabezado: texto{{{texto}}}}}]]] [[[pie: {{{texto}}}}}]]] /// 2: [[[pie: texto]]] /// 3: [[[encabezado: texto]]] /// 4: [[[encabezado: texto]]] [[[pie: texto]]] (Figura 59).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 59. Notación de captura para encabezados y pies de etiquetas. Se muestra un ejemplo de la captura de los encabezados de tres etiquetas en el ejemplar, dos son de determinación (IBUNAM:MEXU:1407723).

Por otro lado, dentro de los encabezados de algunas etiquetas puede encontrarse información de datos primarios y de ocurrencia, que son exclusivos y no están presentes en otra parte de la etiqueta principal o en las secundarias. Por lo que esta información debe ser asignada y capturada a los campos correspondientes, además de aparecer como parte del encabezado o pies según las instrucciones para estos elementos.

De tal modo que si el encabezado o pie de etiqueta contienen datos primarios o de ocurrencia, ya sean taxonómicos, geográficos, de colecta o biológicos, que correspondan a alguno de los campos de la Interfaz de captura, éstos serán extraídos, asignados y capturados como corresponda y siguiendo todas las reglas pertinentes (p. ej., “doble dato”); siempre y cuando en el resto de la etiqueta no se encuentren dichos datos (Figura 60). La única excepción para la asignación de datos geográficos procedentes de encabezados y pies, es aquella que refiere a los nombres de Instituciones, por ejemplo, si el encabezado indica al Herbario Nacional de México, el país “México” no puede ser considerado como tal y asignado al campo de “País”, aun cuando el ejemplar haya sido colectado en México (Figura 61).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

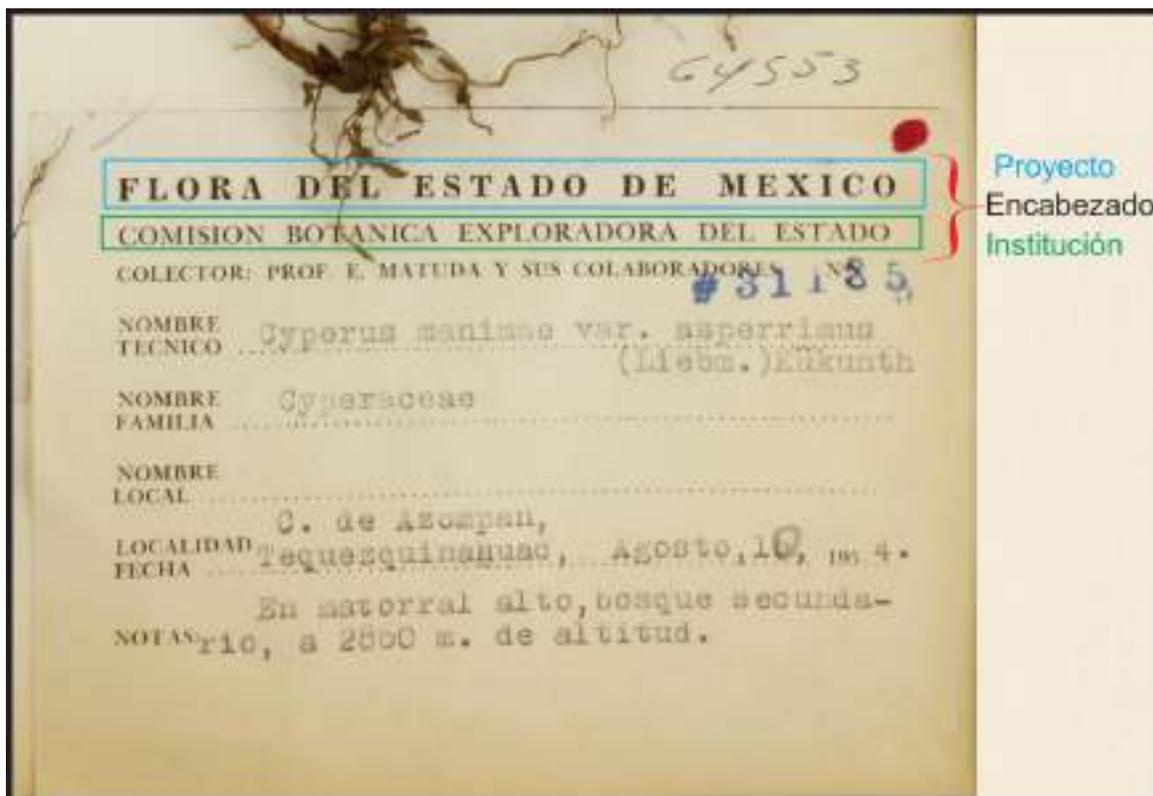


Figura 60. Etiqueta con datos primarios en encabezado. El encabezado contiene el nombre del proyecto (en azul) y la Institución (en verde) (IBUNAM:MEXU:64553). Puede ser asignado el “Estado” (México) a su respectivo campo de la Interfaz de captura.



Figura 61. Datos de encabezados pertenecientes únicamente a Instituciones. Estos datos no pueden ser asignados como datos primarios. a) No se deben extraer datos geográficos ya que especifican los lugares de origen de las instituciones de adscripción de los determinadores, y no



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

corresponden a la geografía del lugar de colecta (IBUNAM:MEXU:1373024); b) No se debe extraer la información señalada en el encabezado (Guadalajara, Jalisco ni México, subrayadas en color rojo) ya que indican la institución donde está depositado el ejemplar. Sin embargo, ya que el encabezado también contiene la leyenda “Plantas de México”, “México” (en color verde) puede capturarse en el campo de país pero debe agregarse la leyenda “confuso” usando el comando de corchete, mientras que existen dos oraciones diferentes con información sobre localidad (Colectas en Sturgeon Bay, Wi. y USSR = Instituto Vavilov, Leningrado, U.R.S.S., ambas en color azul) que deben ser capturadas en el campo de “Localidad” como “doble dato” (IBUNAM:MEXU:938108).

Sin embargo, si la información contenida en el encabezado o pie, no es complementaria o consistente con el resto de los elementos de la etiqueta, dicha información se puede asignar como “doble dato” a los campos correspondientes con el uso del comando de diagonales, e incluso puede usarse el comando de corchetes para incluir la leyenda “confuso” a la información proveniente del encabezado o pie (Figura 62).

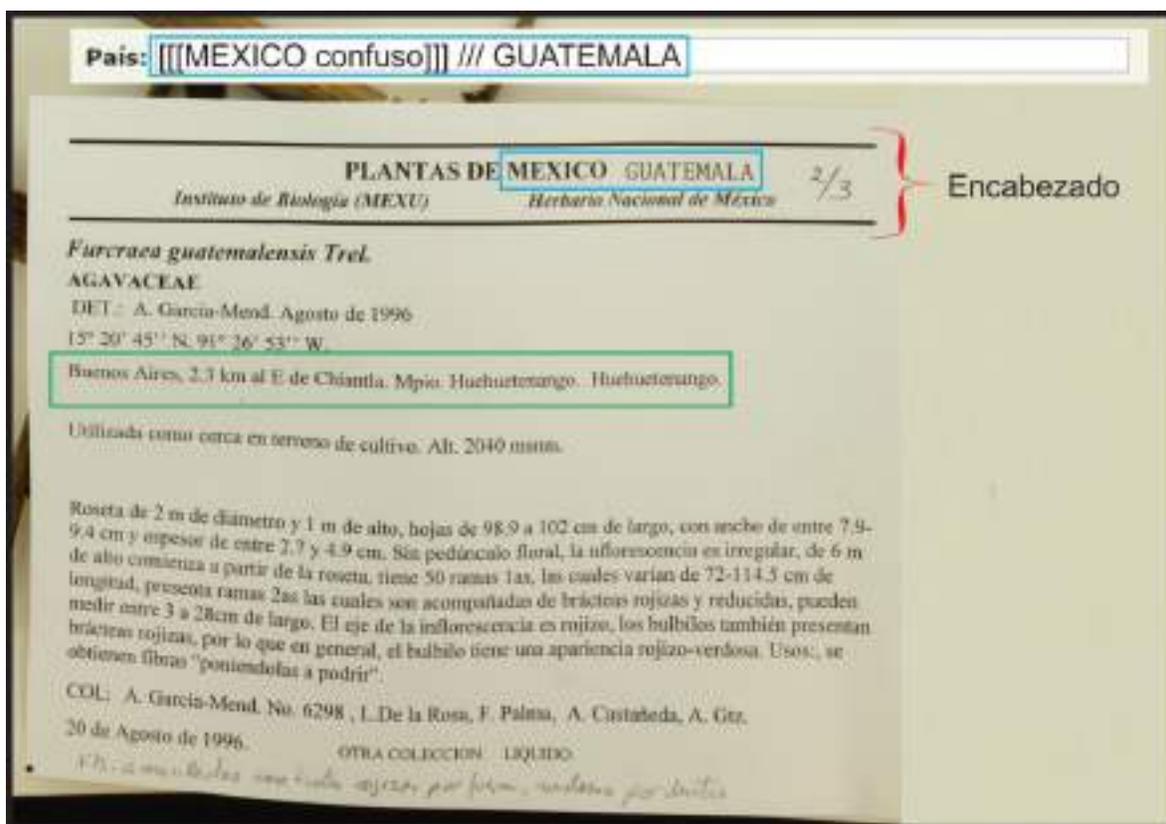


Figura 62. Datos geográficos del encabezado inconsistentes. El encabezado tiene el nombre de dos países diferentes. La localidad (en recuadro verde), da la impresión de pertenecer a “Guatemala”, por ello se debe capturar como doble dato y confuso “México” y “Guatemala” (IBUNAM:MEXU:804292).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.13.12 Sellos

Estos pueden corresponder a etiquetas con sellos que contienen datos exclusivos, que aportan información y no se repiten en el cuerpo de la etiqueta principal (p. ej., determinadores, números de colecta, proyectos, instituciones, otros números de catálogo viejos o de otra institución, etc.).

Si el sello contiene información de datos primarios, que puede ser asignada a su campo correspondiente, debe ser capturada en dicho campo. En cambio si no lo es, debe ser capturado en el campo de “Información adicional”.

Independientemente del campo al que sea asignado la información de un sello, la manera de capturarlos es utilizando la leyenda “Sello” seguido de dos puntos (:) y el texto del sello (p. ej., Sello: texto). Esta instrucción se emplea tan tantas veces como sellos existan asociados al ejemplar, por lo que cada sello es considerado como una partícula independiente y debe ser separada correctamente con el uso del comando de diagonales, cuando sean asignados a “Información adicional”. Además, deben emplearse las instrucciones de los comandos de llaves y corchetes de ser necesario (Figura 63).

Los únicos dos sellos que no se deben capturar son el sello de UNIBIO (que es el logo del Instituto de Biología, por lo general de color verde con o sin la leyenda UNIBIO); así como el sello de MEXU que se presenta en color rojo la mayoría de las veces, en otras azul marino y tiene la leyenda: “Instituto de Biología Herbario Nacional MÉXICO”, y en algunas ocasiones tiene iniciales dentro del sello (Figura 3); sin embargo, no debe confundirse con sellos viejos de MEXU (Figura 63 inciso c).

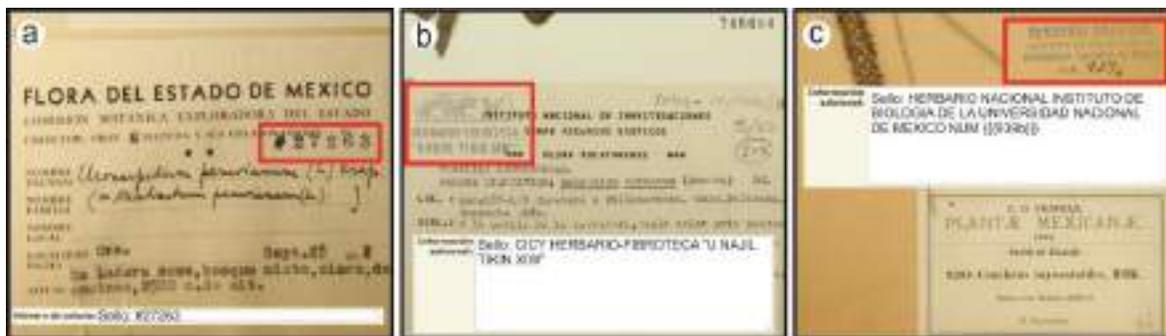


Figura 63. Ejemplos de etiquetas con sellos. Los sellos pueden contener información que no se encuentra en ninguna otra parte de la etiqueta. a) Etiqueta con sello que indica el número de colecta (IBUNAM:MEXU:121692); b) Etiqueta con sello indicando el nombre de un herbario (IBUNAM:MEXU:748684); c) Etiqueta con sello antiguo del herbario MEXU y un número en desuso de catálogo (IBUNAM:MEXU:3952).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 7.2.13.13 Etiquetas con “Información asociada”

Pueden incluir información referente al evento de colecta pero esta es añadida después en una etiqueta distinta de la principal (ver. “Asignación de información asociada”).

### 7.2.13.14 Etiquetas con “Texto asociado”

Pueden incluir desde artículos de revistas científicas o de divulgación, descripciones originales, etc. (ver. “Texto asociado”).

### 7.2.13.15 Etiquetas de “Nota accesoria”

Algunas de estas etiquetas pueden comprender notas sobre toma de muestra o remoción de material, de fotografías, empleados en revisiones taxonómicas, como material de referencia, notas, etc. (ver “Nota accesoria”).

### 7.2.13.16 Anotaciones de tipo curatorial

Estas pueden referir a las siglas de revisiones por parte de taxónomos o personal de las colecciones (p. ej., JLV, VJ, capt., list, etc.). Debe emplearse el comando de llaves de ser necesario.

### 7.2.13.17 Números sin asociación

Algunas etiquetas contienen números sin alguna asociación aparente. Es decir, sin denominadores de campo y sin estar referidos claramente a algún otro valor en particular. Incluso pueden corresponder con números escritos dentro o fuera de las etiquetas o en el cuerpo de la cartulina del montaje. Estos números se deben capturar verbatim y dado que, comúnmente, están escritos a mano debe emplearse el comando de llaves para indicarlo (Figura 64).

Estos números pueden referir a números de clasificaciones taxonómicas de orden dentro de la colección, alguna numeración en desuso, algún código externo a la colección, inclusive números de colecta, pero eso debe ser corroborado con el análisis del resto de los datos de la etiqueta y algunas veces deben compararse con más ejemplares de eventos de colecta similares (incluso de duplicados de otras colecciones). Por lo que no deben asignarse como números de colecta (Ver. “Número de colecta”).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 64. Etiquetas con números sin asociación. a) Existen varios números en la esquina inferior izquierda del ejemplar y otro dentro de la etiqueta junto al colector que no es ni fecha ni número de colecta. (IBUNAM:MEXU:1154419); b) Número fuera de la etiqueta sin aparente asociación (IBUNAM:MEXU:1412537).

### 8. Problemas asociados a imágenes

Esta sección comprende algunos de los problemas comunes asociados al estado de las imágenes. Lo que puede ocasionar que parte de la información de las etiquetas no pueda ser capturada correctamente. Por lo que, se muestran a continuación los problemas principales, la manera de capturar la información y cómo indicar el problema del que se trata.

#### 8.1 Etiqueta doblada

Corresponde a imágenes de ejemplares que contienen alguna etiqueta doblada. Debe capturarse la información legible. El texto que aparezca en el doblez de la etiqueta debe ser sustituido por la leyenda “etiqueta doblada” con el uso del comando de corchetes (p. ej., [[etiqueta doblada]]) (Figura 65). La leyenda puede ser captura en el campo de la información del denominador de campo correspondiente, pero de no ser identificable se debe capturar la leyenda en el campo de “Información adicional”. Además, el problema debe ser reportado a través de la Interfaz de Reportes de problemas en ejemplares.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

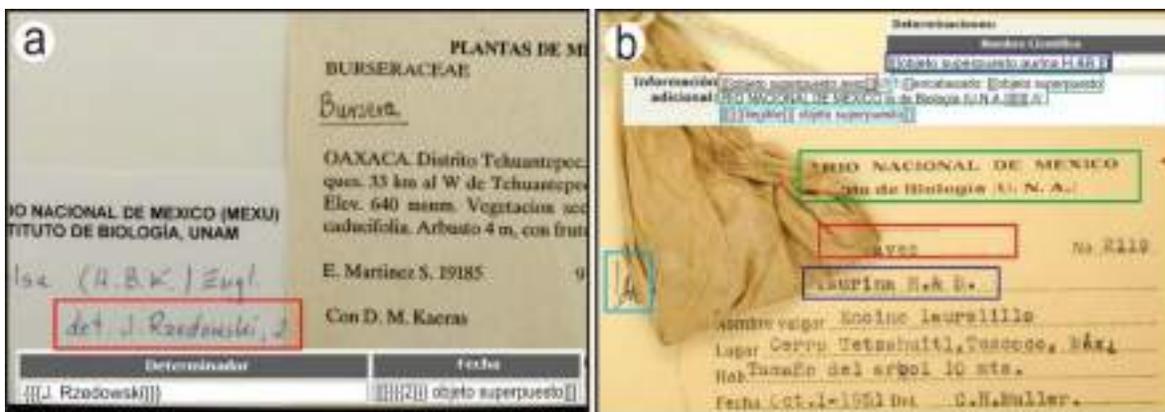
FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 65. Etiquetas dobladas. a) No se puede ver el último dígito del año de colecta (IBUNAM:MEXU:163404); b) Más de la mitad de la etiqueta no puede ser leída. Se ingresa la leyenda en el campo de “Información adicional” porque no es posible saber cuáles y cuántos denominadores de campo están afectados por el doblar de la etiqueta (IBUNAM:MEXU:721230); c) La etiqueta está plegada y dificulta la lectura completa de algunos datos (IBUNAM:MEXU:825137).

### 8.2 Objeto superpuesto

Este problema se presenta cuando existe el traslape de algún objeto sobre algunas de las etiquetas de los ejemplares y no permite la lectura completa de su información. En estos casos se debe capturar la leyenda “objeto superpuesto” con el uso del comando de corchetes, en sustitución de la información que no puede ser visualizada (p. ej., [[[objeto superpuesto]])] (Figura 66). La leyenda puede ser capturada en el campo de la información del denominador de campo correspondiente, pero de no ser identificable se debe capturar la leyenda en el campo de “Información adicional”. Además, se debe de informar sobre este estado en las etiquetas mediante la Interfaz de Reportes de problemas en ejemplares.



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

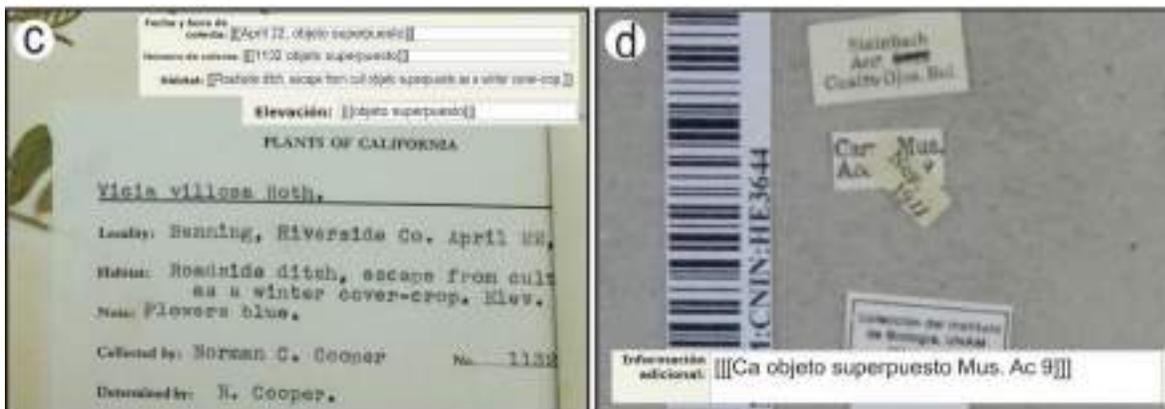
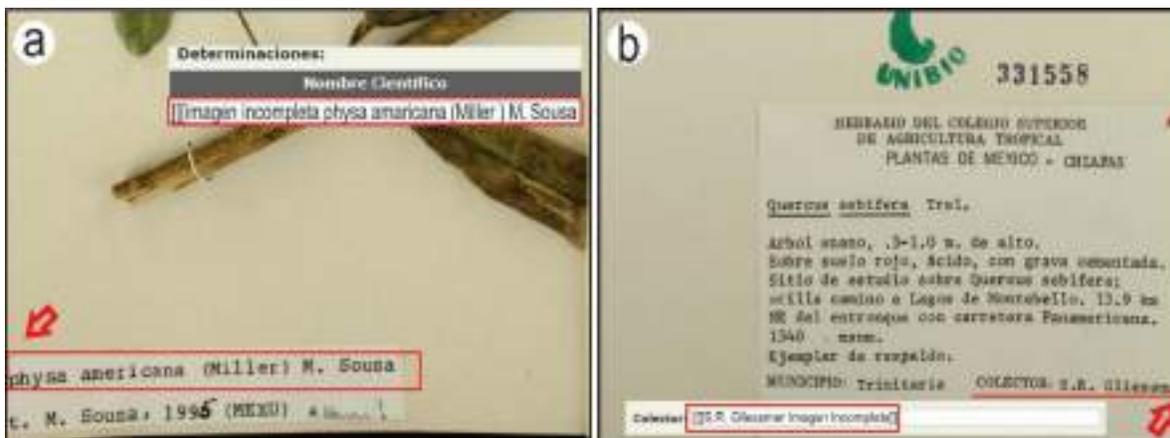


Figura 66. Imágenes con objetos superpuestos sobre las etiquetas. a) La fecha de determinación no se puede visualizar completamente (IBUNAM:MEXU:954069); b) Parte del nombre científico está cubierto, así como otros elementos sin denominador de campo (IBUNAM:MEXU:76904); c) El objeto superpuesto a la derecha de la etiqueta cubre parte de varios campos (IBUNAM:MEXU:457676); d) Una etiqueta cubre a otra y no permite ver qué tipo de información contiene (IBUNAM:CNIN:HE3644).

### 8.3 Imagen incompleta

Aplica cuando la imagen de los ejemplares está incompleta, es decir, uno o más de los bordes de la imagen “corta” parte de la imagen. En estos casos se debe capturar la leyenda “imagen incompleta” con el uso del comando de corchetes, en sustitución de la información que no puede ser visualizada (p. ej., [[[imagen incompleta]]]) (Figura 67). La leyenda puede ser capturada en el campo de la información del denominador de campo correspondiente, pero de no ser identificable se debe capturar la leyenda en el campo de “Información adicional”. Además, se debe de informar sobre este estado en las etiquetas mediante la Interfaz de Reportes de problemas en ejemplares.



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

Sugerencia de cita: DGRU.2020. Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2 (Versión 1). Dirección General de Repositorios Universitarios. SDI-UNAM. México.

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Figura 67. Imágenes incompletas. a) Nombre científico de última determinación incompleto, debe ser *Diphyssa* (IBUNAM:MEXU:645977); b) Apellido de colector incompleto, se lee Gliessmar, cuando debe ser Gliessman (IBUNAM:MEXU:331558).

### 9. Problemas de captura

En esta sección se hace referencia a algunos problemas de captura que se han detectado durante el procesamiento de los datos. Esta mención tiene como objetivo hacer una revisión del tipo de inconsistencias registradas, descritas y categorizadas, lo cual forma parte de la documentación de los procesos de digitalización de los datos de las colecciones del IBUNAM.

De igual forma, se realiza con la intención de incidir en el mejoramiento de la captura de dichos datos, por lo que se hace una descripción breve de los casos señalados, su posible fuente de origen y se establecen las instrucciones para evitar dichos errores.

Se reitera que el presente manual ha de ser usado como una herramienta de prevención de errores, acatando las normas de captura y las sugerencias indicadas, nunca ha de usarse como mecanismo de corrección de errores contenidos en las etiquetas de los ejemplares. Al presentar los casos de inconsistencias se pretende hacer patente que dichos errores se han detectado y los datos se encuentran en proceso de depuración para su uso óptimo.

Los errores de captura se presentan según el término con el cual se han definido para etiquetar dichas inconsistencias durante el proceso de control de calidad (abreviar, adición, asignación, asignación total, completar, digitación, incompleta, procedimiento y omisión). Están listados en orden alfabético, para facilitar su localización, sin embargo este orden no refleja la gravedad de los errores. Cualquiera de los errores puede ser tan grave como otro dependiendo del campo en que se presente y la forma en que altere la información. Sin embargo, “Omisión” y “Asignación total” resultan ser los más graves, debido a la dificultad para su detección.

En algunos casos se muestra la forma en que fue capturada la información y porque es considerado un error de captura. La manera correcta de capturar la información ya se ha explicado en la sección 7 de esta guía.

#### 9.1 Abreviar

Cuando se abrevian nombres propios o palabras presentes en los enunciados de la descripción de localidades, notas de campo, etc. (Figura 68).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

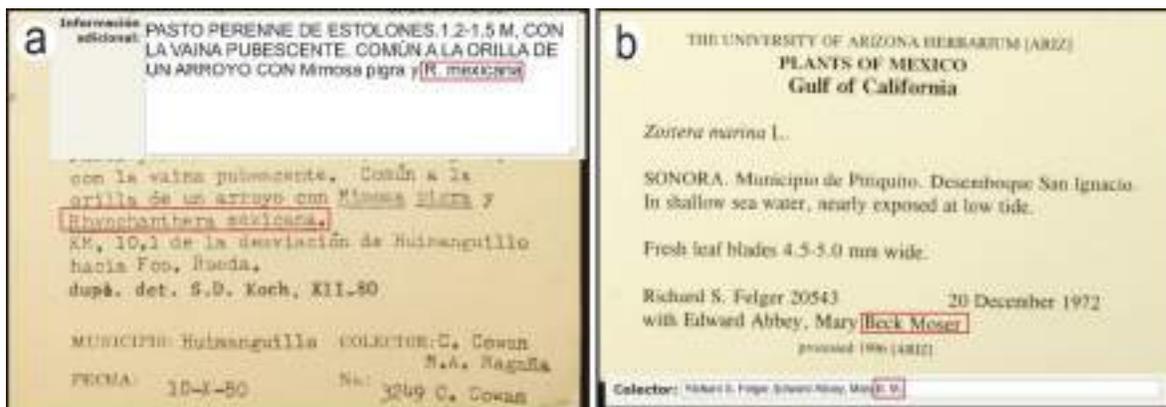


Figura 68. Ejemplos de capturas con abreviaturas. a) En la segunda especie asociada (*Rhynchanthera mexicana*) fue abreviado el género (IBUNAM:MEXU:327113); b) Se abrevió el nombre del último colector (IBUNAM:MEXU:777907).

### 9.2 Adición

Se agregan letras, palabras, números o símbolos que no están presentes en las etiquetas. Este tipo de error es mayor a la añadidura producida en los “Errores de digitación” (Figura 69).

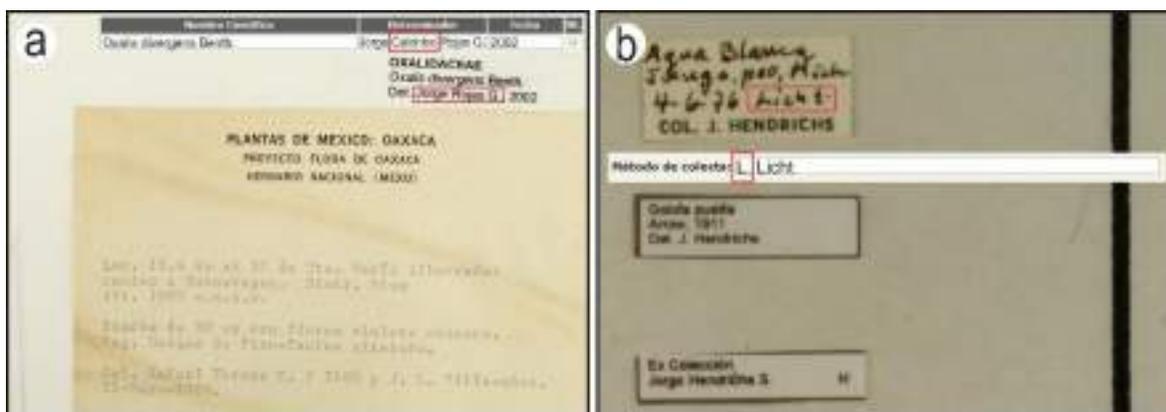


Figura 69. Errores de adición. a) Se agregó el apellido “Calónico” al nombre del determinador, lo que crea un nombre inexistente (IBUNAM:MEXU:1132839); b) Se agregó “L.” al campo “Método de colecta” (IBUNAM:CNIN:CO33355).

### 9.3 Asignación

Esta ocurre cuando la información no corresponde a la naturaleza del campo o el dato asignado es discordante entre el presente en la etiqueta y el capturado, a pesar de tener un denominador de campo evidente (Figura 70).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Un ejemplo común de asignación es la captura de información correspondiente a campos definidos en la Interfaz de captura en el campo de “Información adicional”. Por ejemplo, términos que están asociados directamente al nombre científico, que siguen la regla de “Asignación de información asociada”.



Figura 70. Ejemplos de errores de asignación. a) La partícula “ined.” no fue capturada junto al nombre científico (IBUNAM:MEXU:433843); b) La partícula “comb nov. ined.” fue capturada en el campo de “Información adicional” (IBUNAM:MEXU:666975); c) El texto de “hábitat” no fue capturado en el campo correcto a pesar de que tiene un denominador de campo evidente (IBUNAM:MEXU:1329623).

### 9.4 Asignación total

Ocurre cuando se captura información que no se encuentra en la etiqueta, aun cuando el texto ingresado sea correcto o consistente con el resto de la información de la etiqueta (Figura 71). Estos casos sobrepasan a los errores de “Digitación” y “Adición”, ya que cambia por completo el sentido de la información.

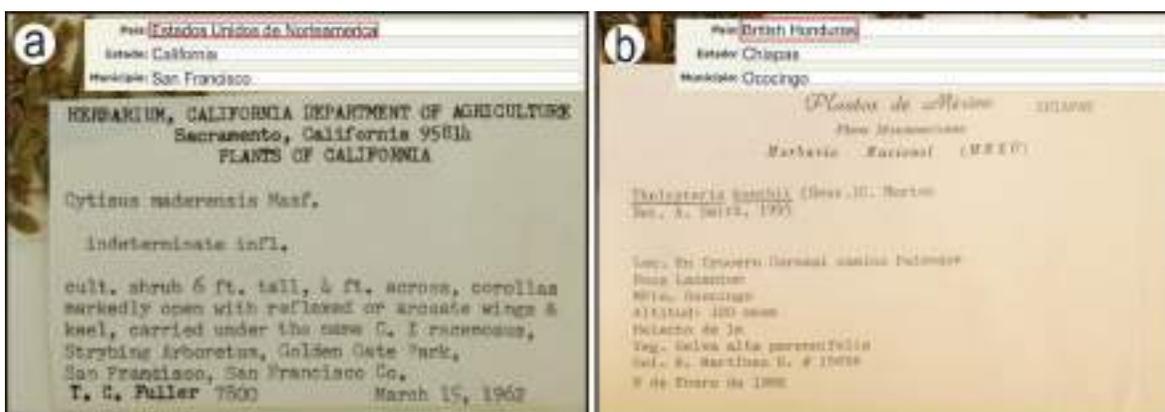


Figura 71. Errores de asignación total. a) Asignación total de país, aun cuando sí es de Estados Unidos de América, no aparece referido en la etiqueta y además está mal escrito (IBUNAM:MEXU:854919); b) Asignación total de país, además de que no aparece en la etiqueta, no es un nombre válido de país (IBUNAM:MEXU:680646).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 9.5 Completar

Es el error contrario a “Abreviar”. Si la información viene abreviada en la etiqueta, es completada en la captura. En ocasiones, incluso, se completa de manera errónea (Figura 72).

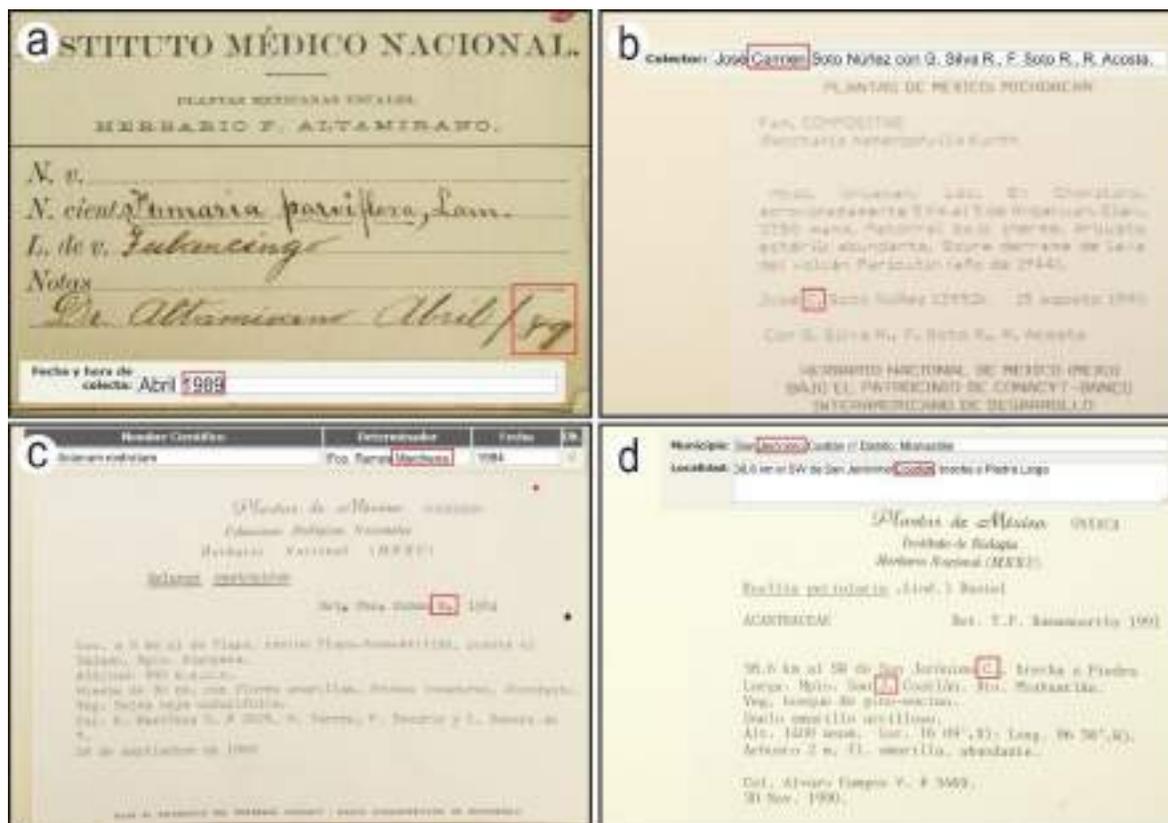


Figura 72. Errores de captura en los que se completa la información. a) Se completó el año de colecta (89 a 1989). El colector murió en 1908 y la institución dejó de existir en 1915, por lo que el año correcto es 1889 (IBUNAM:MEXU:53696); b) Se completó “C.” como “Carmen”, aun así el nombre es correcto (IBUNAM:MEXU:819581); c) Se completó el segundo apellido, a pesar de ello es correcto (IBUNAM:MEXU:369009); d) El nombre del municipio y la localidad son el mismo aunque es incorrecto completarlos (IBUNAM:MEXU:553842).

### 9.6 Digitación

Aplica a aquellos casos donde se han adicionado, cambiado u omitido caracteres (letras, números, símbolos, etc.) en la captura. Este tipo de error incluye errores de ortografía que no sean de origen (es decir que no sean verbatim). Generalmente las inconsistencias causadas por estos errores son menores a los generados por errores de “Adición” y “Asignación total” (Figura 73).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020



Figura 73. Errores de digitación en captura. a) El apellido del colector tiene una letra adicional (IBUNAM:MEXU:1008360); b) La inicial del colector asociado cambia de “J.” por “A.” (IBUNAM:MEXU:1152049); c) El apellido del primer colector se escribió mal, mientras que en el caso del tercero el error ortográfico genera un apellido completamente diferente (IBUNAM:MEXU:631706).

### 9.7 Incompleta

Cuando no se captura la información completa del registro, es decir que se trunca en algún punto (p. ej., un enunciado de 10 palabras, de las que solo se capturan cinco). Este error es mayor al producido por el error de “Digitación”; y junto con “Omisión” son difíciles de detectar y pueden generar errores más graves en los procesos posteriores a la captura (p. ej., podrían crearse variantes de nombres de colectores y determinadores incorrectos o inexistentes) (Figura 74).

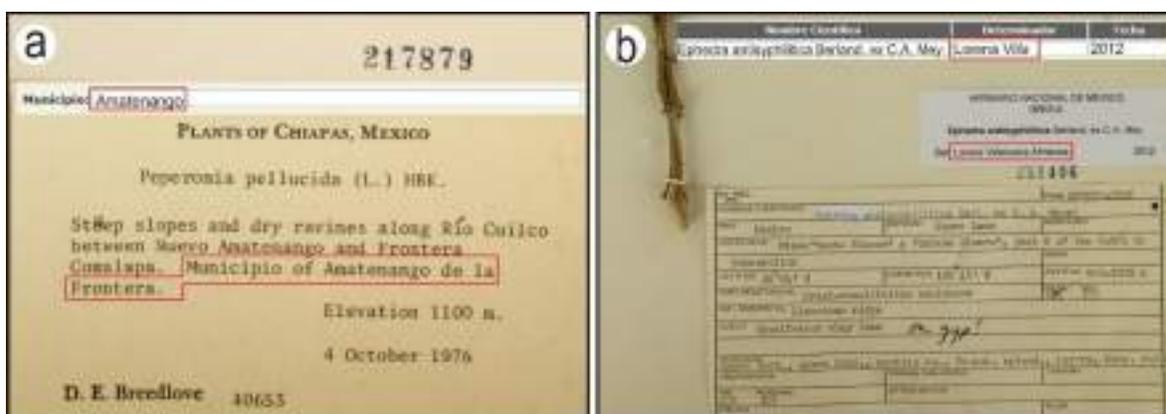


Figura 74. Ejemplos de captura incompleta. a) Sólo fue capturado “Amatenango” en el campo de “Municipio”, sin embargo para el estado de Chiapas existen dos municipios cuyo nombre contiene la palabra “Amatenango” (Amatenango de la Frontera y Amatenango del Valle), esto ocasiona que no pueda ser validado correctamente el municipio (IBUNAM:MEXU:217879); b) Se capturó sólo parte de los apellidos del determinador, lo cual genera un nombre incorrecto (IBUNAM:MEXU:251406).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 9.8 Procedimiento

Son errores producidos por la captura de la información en formatos distintos a los presentados en las etiquetas, es decir, captura no verbatim (Figura 75). También, pueden ser considerados aquellos en los que se omiten o se incumplen las reglas de captura establecidas (generales o específicas), según sea el caso.

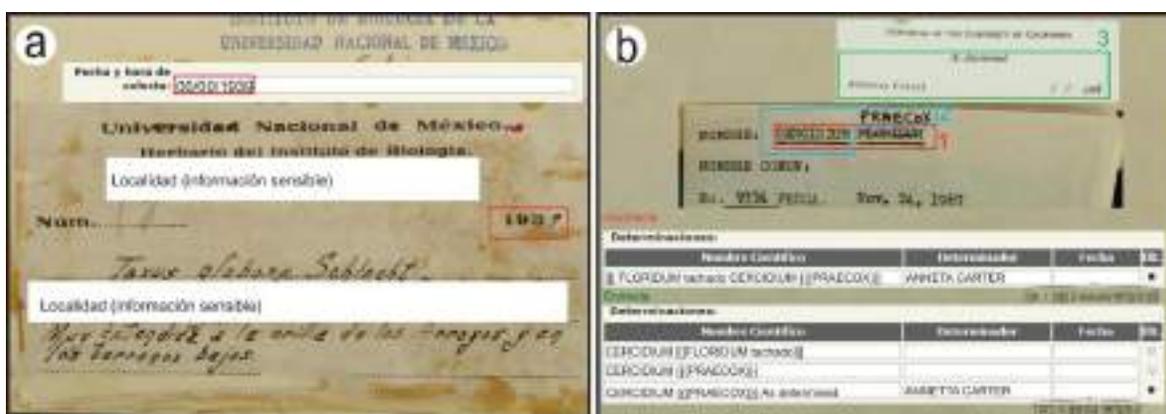
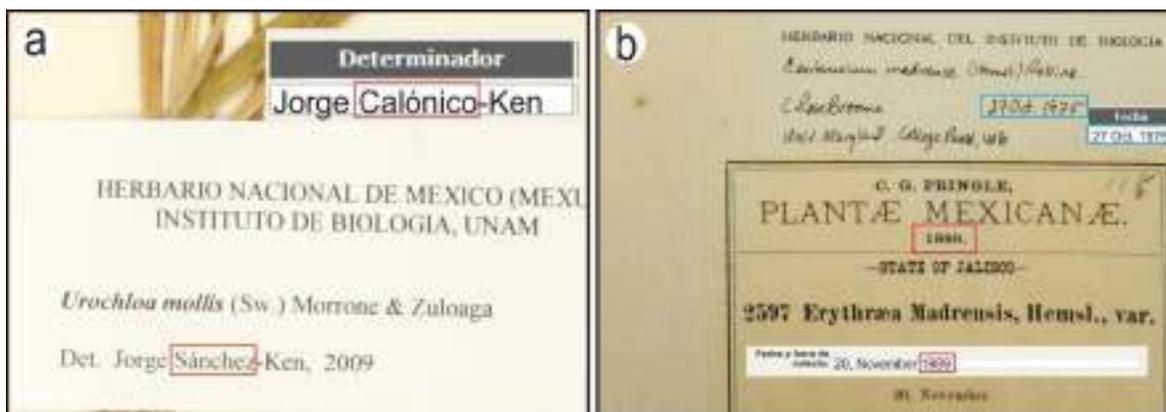


Figura 75. Capturas con errores de procedimiento. a) No fue respetado el formato original de la fecha de colecta (IBUNAM:MEXU:2451); b) No fueron asignados correctamente los eventos de determinación y las reglas de captura fueron combinadas incorrectamente (IBUNAM:MEXU:180248).

Dentro de los problemas ocasionados por una mala captura, pueden presentarse una serie de combinaciones de errores que generan inconsistencias mayores, lo cual hace más complicada su detección. En algunos casos estas inconsistencias pueden hacer pensar que se trata de un error de origen, pero en realidad son errores de captura, tal es el caso de no capturar determinada información (“Omisión”) y añadir otra inexistente en la etiqueta (“Asignación total”) (Figura 76).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

Figura 76. Ejemplos de capturas inconsistentes. a) Se omitió el apellido del determinador y fue sustituido por el de otro, lo que genera un nombre inexistente (IBUNAM:MEXU:1081034); b) Las fechas de colecta y determinación capturadas son inconsistentes, ya que la de colecta es posterior a la de determinación, lo que es ilógico. Sin embargo, esto no es un error de origen sino un error de captura en la fecha de colecta, debe ser 1889 en lugar de 1989 (IBUNAM:MEXU:24939).

Dichas inconsistencias en la captura pueden verse agravadas por ambigüedades inherentes a la misma información de las etiquetas. Por lo que si la información es incierta, dudosa o confusa debido a la vaguedad en su presentación, pueden presentarse combinaciones de errores adicionales, que pudieron haberse evitado de mantener la captura verbatim (Figura 77). De este tipo de ambigüedades no incluyen las causadas por el estado físico de los ejemplares, imágenes y sus etiquetas (p. ej., imagen incompleta, objeto superpuesto, etiqueta dañada, etc.).



Figura 77. Ejemplos de información ambigua en etiquetas. a) La traducción del denominador “Reçule” es recibido, no colectado. Por lo que no representa la fecha de colecta del evento. Sin embargo, al no existir otra fecha y posiblemente desconocer el significado del denominador, poco común, fue capturado como fecha de colecta (IBUNAM:MEXU:1300384); b, c y d) (IBUNAM:MEXU:1389487, IBUNAM:MEXU:1397466, IBUNAM:MEXU:1414151, respectivamente).



CC-BY-4.0 Internacional. <https://creativecommons.org/licenses/by/4.0/legalcode.es>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

La asignación de la información geográfica (división administrativa) es diferente aun siendo el mismo formato de etiqueta. En este caso existen tres niveles de división administrativa, pero no están indicados claramente y dado que el tercer nivel de la etiqueta “Municipio” tiene el mismo denominador de campo que el segundo nivel de la Interfaz, esto complica la asignación correcta de la información. Adicionalmente en el inciso (b) se agregó erróneamente “Depto” en el campo de “Estado” y se abrevió “Provincia” en el campo de “Municipio”, aunque se asignó correctamente según los tres niveles de división administrativa. La opción más recomendable es el inciso (c), considerando las reglas de captura de “Denominador de campo” y de “Asignación y atomización de la información”.

Otros errores comunes de procedimiento incluyen la mala atomización. En estos casos puede tratarse de aquellos ejemplares que al ser capturados no se separan correctamente los datos, con el uso del comando de diagonales en el caso del campo de “Información adicional”. Así mismo, pueden incluirse errores en los que no se usan correctamente los separadores propios de la información de la etiqueta (p. ej., signos de puntuación o su ausencia) o el uso del comando de diagonales en sustitución de la puntuación de la información de la etiqueta (Figura 78).



Figura 78. Errores de mala atomización de la información. a) Se utilizó el comando de diagonales en sustitución los signos de puntuación del texto de “Habitat” (IBUNAM:MEXU:1327227); b) Los “guiones” fueron omitidos como separadores y fueron sustituidos por espacios (IBUNAM:MEXU:276319).

La importancia de la captura verbatim radica, entonces, en el hecho de que si existe una inconsistencia en la información capturada de los ejemplares es debido a un error de origen. Es decir, que no se trata de un problema de captura, sin embargo, estos registros se califican de manera específica para su procesamiento en el control de calidad. Por lo que la captura verbatim permite detectar más fácilmente estos problemas, sin tener que extender la búsqueda a más ejemplares o verificar la captura directamente con la imagen del ejemplar (Figura 79).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

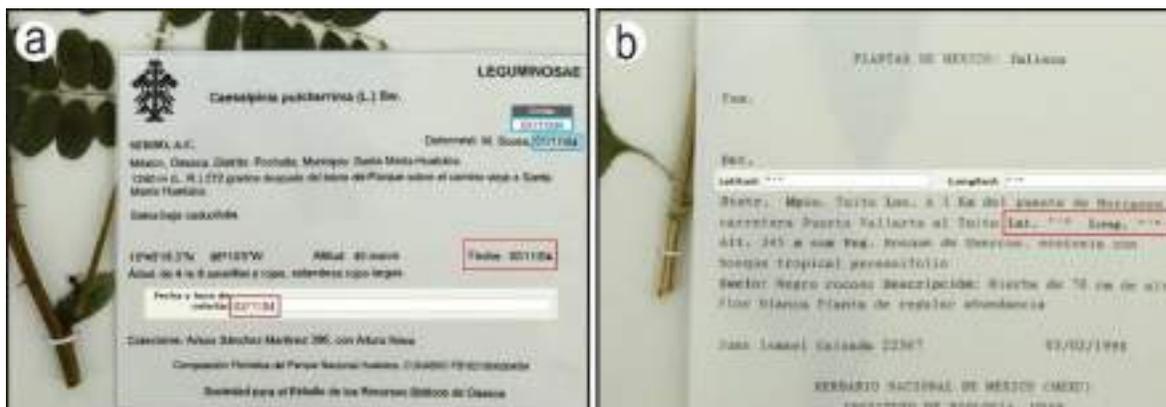


Figura 79. Captura verbatim. a) La fecha de colecta del ejemplar es posterior a la determinación, lo que es inconsistente. Sin embargo, es un error de origen y fue capturado correctamente. Es posible que las fechas estén invertidas o puede ser que alguna sea incorrecta, por lo que necesita ser investigado o revisado por el curador de la colección (IBUNAM:MEXU:1206691); b) La captura de las coordenadas es verbatim. Igualmente corresponde a un error de origen, en este caso puede interpretarse como ausencia de datos. (IBUNAM:MEXU:1087633).

### 9.9 Omisión

Es un tipo de error en el que la información contenida en la etiqueta no fue capturada. La omisión es un error muy grave (superior a “Incompleta”) y aún más difícil de detectar. Puede incluir desde palabras, frases, enunciados, párrafos, hasta etiquetas completas (Figura 80). Esto incluye la omisión de elementos estructurales de las etiquetas (encabezados y pies) y sellos.

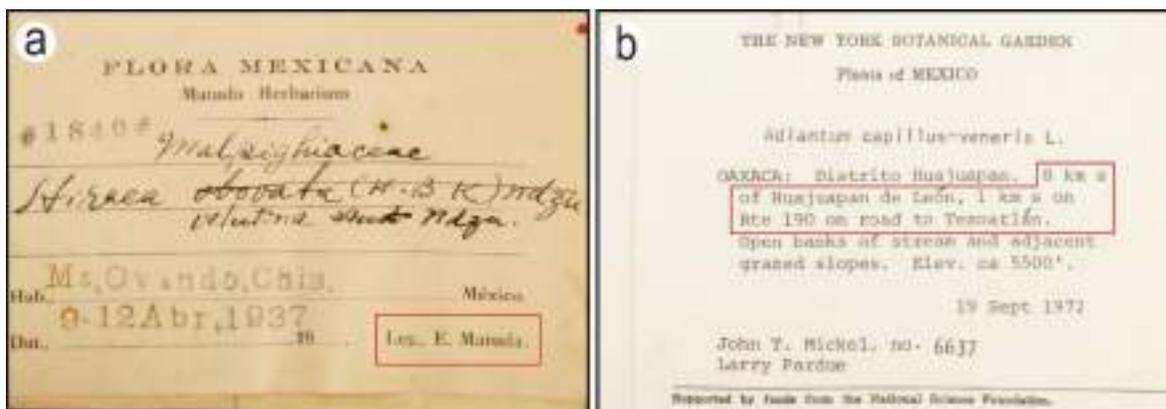


Figura 80. Errores de omisión en captura. a) No se capturó el nombre del colector (IBUNAM:MEXU:77922); b) Fue omitida la localidad (IBUNAM:MEXU:873649).

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 10. Bibliografía

Chapman, A. D. (2005). *Principles of Data Quality*. Global Biodiversity Information Facility. Recuperado de <https://doi.org/10.15468/doc.jrgg-a190>

Chapman, A. D. & Busby J. R. (1994). Linking plant species information to continental biodiversity inventory, climate and environmental monitoring. En Miller, R. I. (Ed.), *Mapping the Diversity of Nature* (pp. 177-195). London: Chapman and Hall.

Dalcin, E. C. (2004). *Data Quality Concepts and Techniques Applied to Taxonomic Databases* (Tesis doctoral). Escuela de Ciencias Biológicas, Universidad de Southampton, Reino Unido.

Gernandt, D. S., Sánchez-Cordero, V., Samper, U. M., Giménez, O. J., & Salazar, G. (2014). Digitalización del Herbario Nacional de México: avances y retos del futuro. *Revista Digital Universitaria*, 15, 1–13. Recuperado de <http://www.revista.unam.mx/vol.15/num4/art30/>

Dasu, T., Vesonder, G. T. & Wright, J. R. (2003). Data quality through knowledge engineering. En L. Gettor (Presidencia), *KDD '03 Proceedings of the Ninth International Conference on Knowledge Discovery and Data Mining*. Conferencia llevada a cabo en el congreso de la Association for Computing Machinery, Washington, D.C., U.S.A. doi: <http://doi.acm.org/10.1145/956750.956844>

English, L. P. (1999). *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*. New York: John Wiley & Sons, Inc.

Fürber, C. & Hepp, M. (2011). Towards vocabulary data quality management semantic web architectures. En T. Risch (Presidencia), *Proceedings of the 1st International Workshop on Linked Web Data Management*. Conferencia llevada a cabo en la EDBT/ICDT 2011 Joint Conference, Uppsala, Sweden. doi: <https://dl.acm.org/citation.cfm?doid=1966901.1966903>

Gartner. (2007). *'Dirty Data' is a Business Problem, Not an IT Problem, Says Gartner*. Sydney, Australia: Gartner Newsroom. Recuperado de <http://www.gartner.com/newsroom/id/501733>



## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

---

Maletic, J. I. & Marcus A. (2000). Data Cleansing: Beyond Integrity Analysis. En Y. W. Lee & S. E. Madnick (Co-presidencia), *Proceedings of the Conference on Information Quality*. Conferencia llevada a cabo en la IQ2000, Massachusetts Institute of Technology, Boston. doi: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.37.5212>

Redman, T. C. (2001). *Data Quality: The Field Guide*. Boston, Massachusetts: Digital Press.

Singh, R. & Singh. K. (2010). A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing. *International Journal of Computer Science Issues*, 7,(3), 41-50.

Tegelberg, R., Haapala, H., Mononen, T., Pajari, M., & Saarenmaa, H. (2012). The development of a digitising service centre for natural history collections. *ZooKeys*, 209, 75-86. doi: <https://doi.org/10.3897/zookeys.209.3119>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### 11. Anexos

#### Anexo A. Lista de situaciones problemáticas que deben reportarse en la Interfaz de Reportes de problemas en ejemplares.

##### Recepción de imágenes:

- Ausencia de folio o número de catálogo, así como de los sellos de UNIBIO y los pertinentes para cada colección
- Doble folio (cuando sólo existe un evento de colecta)
- Montajes donde se encuentra más de un ejemplar de la misma o diferente especie, que corresponden a distintos eventos de colecta, asociados al mismo número de identificación
- Montajes que son fotografías de un ejemplar, es decir, que no son herborizado
- Nitidez y legibilidad de las imágenes de los ejemplares. Incluye aquellas dañadas durante la migración
- Tachaduras o enmendaduras en los números de folio o de catálogo

##### Reglas de captura:

- Confusa
- Dañada
- Error de origen
- Etiqueta doblada
- Extranjera
- Ilegible
- Imagen incompleta
- Múltiples (inconsistentes)
- Objeto superpuesto
- Opciones múltiples, de llenado o marcado (sin opción)
- Probable error

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

### Anexo B. Reglas asociadas a los comandos (corchetes, llaves y separadores) en la Interfaz de captura.

Comando	Descripción	Tecla	Reglas asociadas
Corchetes ([[[ ]]])	triplete de corchetes de apertura y de cierre, dentro del cual se ha de escribir la partícula en cuestión	F9	<ul style="list-style-type: none"> <li>• Adicional al evento</li> <li>• Confusa</li> <li>• Dañada</li> <li>• Encabezados y pies de etiquetas principales y secundarias</li> <li>• Enmendadura</li> <li>• Etiqueta doblada</li> <li>• Extranjera</li> <li>• Imagen incompleta</li> <li>• Múltiples</li> <li>• Nota Accesoría</li> <li>• Objeto superpuesto</li> <li>• Opciones múltiples, de llenado o marcado</li> <li>• Texto asociado</li> </ul>
Diagonales (///)	triplete de diagonales, y un espacio simple antes y después del triplete	F4	<ul style="list-style-type: none"> <li>• Atomización (sólo en "Información Adicional")</li> <li>• Doble dato</li> <li>• Ortotipografía y correcciones</li> <li>• Varias Familias</li> </ul>
Llaves ({{{ }})	triplete de llaves de apertura y de cierre, dentro del cual se ha de escribir la partícula en cuestión	F1	<ul style="list-style-type: none"> <li>• Escrita a mano</li> </ul>

## Guía para la unificación de criterios de captura de datos en estándar Darwin Core 2

FECHA DE ELABORACIÓN: 6 OCTUBRE 2020

---

### Anexo C. Buenas prácticas en la captura de los datos.

- Revisar los ejemplares e imágenes a capturar
- Utilizar adecuadamente la Interfaz de Reportes de problemas en ejemplares
- No añadir saltos de renglón (*enter*) separar datos con espacio simple (barra espaciadora)
- Extraer la información de manera sistemática y ordenada, etiqueta por etiqueta y asignar la información a los campos correspondientes de la Interfaz de captura, evitando “rellenar” dichos campos en el orden en el que se presentan
- Capturar toda la información existente dentro y fuera de etiquetas
- Seguir las reglas de captura establecidas en este manual
- En el caso de los eventos de determinación capturarlos siguiendo el orden cronológico en el que aparecen
- No interpretar datos ilegibles
- No asignar signos de interrogación sustituyendo los datos ilegibles
- No abreviar datos que en la etiqueta se encuentran *in extenso*
- No completar datos que en la etiqueta se encuentran abreviados
- Utilizar adecuadamente las opciones de repetir anterior
- Revisar los registros capturados para corroborar los datos