

Darwin Test

Versión 3.1

Manual de uso
Unidad de Coordinación GBIF.ES
2011

ÍNDICE

1. INTRODUCCIÓN.....	2
1.1. Versiones y créditos	4
2. GUÍA DE INSTALACIÓN Y USO	7
2.1. Descarga del programa	7
2.2. Validar los datos	8
2.3. Comprobar los datos	11
2.4. Corregir los datos	13
2.5. Traspasar los datos corregidos a la tabla darwincore	14
3. DETECTAR CARACTERES ASCII ANÓMALOS	15
4. CONFIGURAR Y PERSONALIZAR LA APLICACIÓN.....	17
4.1. Configurar la tabla de Validaciones	17
4.2. Elegir el grado de generalización de las coordenadas	22
4.3. Vincular la tabla de Species2000 o Archivos de autoridad taxonómica	22
5. CONTROL DE MENSAJES DE ERROR DURANTE LA VALIDACIÓN	24
6. CONVERTIR COORDENADAS A COORDENADAS DECIMALES	26
7. DIFUMINAR O ELIMINAR COORDENADAS DE DARWINCORE	28
8. ELIMINAR REGISTROS DE DARWINCORE	32
9. EXPORTACIÓN A DARWIN CORE ARCHIVE.....	33
9.1. Metadatos de la colección	34
9.2. Creación del archivo Darwin Core Archive	41
10. IMPORTAR ARCHIVO DARWIN CORE ARCHIVE.....	42
11. ÍNDICE DE CALIDAD APARENTE.....	43
APÉNDICE I: TABLAS.....	45
APÉNDICE II: COMPACTAR Y REPARAR	47

1. INTRODUCCIÓN

DARWIN TEST es una aplicación para la validación y chequeo de los registros presentados en tablas **DARWINCOREV2** y **DARWINCORE14** (formato estándar Darwin Core versión 1.2 y Darwin Core 1.4 para el intercambio de información de especímenes y observaciones de colecciones de historia natural, <http://rs.tdwg.org/dwc/terms/history/versions/index.htm>).

Se recomienda la validación de los datos en las tablas **DARWINCORE** utilizando **DARWIN TEST** antes de poner los datos en el servidor para su visualización en la web de la red de GBIF (<http://www.gbif.es/>)

Los chequeos que realiza son de varios tipos:

TIPO DE ERROR	TIPO DE VALIDACIÓN	CAMPO QUE VALIDA
Errores de omisión	Información ausente	Número de catálogo
		Nombre de la institución
		Nombre de la colección
		BasicOfRecord
Errores tipográficos	Ajuste a reglas de nomenclatura	Género
		Especie
		Subespecie
Errores de Convención	Formato de campo	Coordenadas geográficas
		Campos numéricos
	Errores de ajuste a valores predeterminados contenidos en tablas de la aplicación	Reinos
		Continentes
Ajuste a estándares ISO contenidos en tablas de la aplicación	Países	
	Provincias	
Ajuste a valores definidos en la propia consulta	Autor del nombre científico	

		BasicOfRecord
Errores de Congruencia	Ajuste a valores lógicos	Fechas
		Máxima y mínima altitud
		Máxima y mínima profundidad
	Ajustes a valores de bounding box de países y provincias	Coordenadas geográficas
Chequeo contra bases de datos externas (Species 2000)	Género	
	Especie	
	Subespecie	

Tabla 1. – Chequeos que realiza Darwin Test.

DARWIN TEST está basado en *MS-Access*, funciona con sistema operativo Windows y MS-Office XP instalado. **Organización:**

- Consta de una única base de datos: **Darwin_test.mdb**.
- Las tablas de estándares y las tablas generadas por la aplicación van en Mayúsculas y son accesibles y modificables por los usuarios.
- Las consultas de chequeo y validación pueden ser modificadas por el usuario según sus necesidades.

La aplicación en su versión completa está disponible en la zona de descarga de la página web de GBIF http://www.gbif.es/darwin_test/Darwin_Test.php, junto con el 'Manual de Usuario' en formato *pdf*. Puede descargarse tanto el programa ejecutable (**Darwin_Test.mde**) como el programa fuente (**Darwin_test.mdb**: desde la página de código abierto: <http://sourceforge.net/projects/darwintest14/>.)

Una vez descargada la aplicación (ver el punto [2. Guía de Instalación](#)), está lista para su uso, y puede también cambiar su configuración si lo desea. En todo caso, puede consultar la configuración en el punto [4. Configurar y personalizar la aplicación](#)).

CONVENCIONES DE ESTILO Y SÍMBOLOS UTILIZADOS EN ESTE MANUAL:

ESTILO:

- Los nombres de bases de datos van en minúscula y negrita: **Darwin_Test.mde**.
- El nombre de la aplicación va en mayúsculas: DARWIN TEST.
- Los nombres de tablas van en mayúscula y negrita: **ISO3166, REINOS, CONTINENTES**, etc.
- Los nombres de campos van en mayúsculas: GENUS, SPECIES, SCIENTIFICNAME, etc.
- Los botones van en negrita con fondo gris: **Configurar**.

SÍMBOLOS:





NOMBRE	ICONO	ACCIÓN
Lápiz nuevo registro		Cuando aparece este lápiz en el margen izquierdo de la ventana estamos modificando el registro que está a la vista.
Navegación		Permite navegar a través de las rutas (directorios) para elegir la ruta y archivos necesarios para la aplicación.
Nuevo registro		Añade un nuevo elemento (registro): da de alta un registro.

Tabla 2. - Simbología usada en este manual

IMPORTANTE. Cada vez que veamos el icono  en el margen izquierdo de la ventana significa que estamos editando el registro y se debe prestar atención para no cambiar o perder algún dato de forma no consciente, y pulsar la tecla **Escape (Esc)** en caso de estar modificando algún dato sin querer.

1.1. VERSIONES Y CRÉDITOS

Versión 1.0: Primera versión operativa en SourceForge.net.

Versión 1.1: Desarrollada y programada por Isabel Ortega Maqueda, Unidad de coordinación de GBIF España, Septiembre 2007. Primera versión pública.

Versión 1.2: Desarrollada por Isabel Ortega Maqueda, Unidad de coordinación de GBIF España, Abril 2008.

Versión 1.3: Desarrollada por Isabel Ortega Maqueda, Unidad de coordinación de GBIF España, Diciembre 2008.

Versión 1.4: Desarrollada por Isabel Ortega Maqueda y M^a Carmen Lujano Bermúdez. Unidad de coordinación de GBIF España, Octubre 2009.

Versión 2.0: Desarrollada por Isabel Ortega Maqueda, M^a Carmen Lujano Bermúdez y Jesús Rodríguez Escribano. Unidad de coordinación de GBIF España, Febrero 2010.

Versión 2.1: Desarrollada por Isabel Ortega Maqueda, M^a Carmen Lujano Bermúdez y Jesús Rodríguez Escribano. Unidad de coordinación de GBIF España, Enero 2011.

Versión 3.0: Desarrollada por Isabel Ortega Maqueda, M^a Carmen Lujano Bermúdez y Jesús Rodríguez Escribano. Unidad de coordinación de GBIF España, Agosto 2011.

Versión 3.1: Desarrollada por Isabel Ortega Maqueda, M^a Carmen Lujano Bermúdez y Jesús Rodríguez Escribano. Unidad de coordinación de GBIF España, Noviembre 2011.

Características

- Permite **validar** y **chequear** los campos de las tablas **DARWINCOREV2** o **DARWINCOREV14** (formato estándar Darwin Core versión 1.2 y Darwin Core 1.4) con consultas predeterminadas configurables.
- Permite la **activación/desactivación** de las consultas existentes.
- Permite la **creación de nuevas consultas** de validación adaptadas a las necesidades de cada usuario.
- Ofrece la posibilidad de **corregir los errores** detectados en los datos desde los propios formularios de validación. Dichos cambios pueden integrarse de forma permanente en la

tabla Darwincorev2 o 1.4 de origen o reflejarse en una tabla intermedia (Darwincorev2Validada/ Darwincore14Validada).

- Ofrece la posibilidad de **importar archivos con formato Darwin Core Archive** a las tablas **DARWINCOREV2** y **DARWINCOREV14** para realizar la validación de su contenido.
- Así mismo, permite **exportar** las tablas **DARWINCOREV2** y **DARWINCOREV14** al formato **Darwin Core Archive**.
- Permite el chequeo de los nombres científicos con la **Tabla de Nombres SCIENTIFIC_NAMES** del consorcio [The Catalogue Of Life: JTIS y Species 2000](#), con la **tabla Archivos de Autoridad Taxonómica (AAT)** del [SIB de Colombia](#), o con otra base de datos con nombres científicos que cumpla unos requisitos mínimos.
- Permite **convertir coordenadas UTM, MGRS o coordenadas sexagesimales en coordenadas geográficas en grados decimales**, utilizadas en las tablas **DARWINCOREV2** y **DARWINCOREV14**.
- Chequea la existencia de **caracteres ASCII anómalos** y los reemplaza por caracteres válidos que puede definir el usuario.
- Permite crear un **índice de calidad aparente (ICA)** para el seguimiento de la mejora de la calidad de los datos.
- Permite la **creación y gestión de filtros** de usuario para eliminar registros o generalizar coordenadas selectivamente de las tablas **DARWINCOREV2** y **DARWINCOREV14**.
- **Código fuente** accesible desde la **página del proyecto en Sourceforge.net** bajo licencia [Creative Commons](#).

Otros Créditos

Las funciones que realizan la conversión provienen del programa MANDRAS de la antigua DMA, ahora NIMA (National Imagery and Mapping Agency--E.E.U.U., <http://www.nima.mil/GandG/metadata/dt-md.html>), cuyas funciones fueron traducidas de FORTRAN a PowerBASIC por Yvan Barbier y Pierre Rasmont (Laboratoire Zoologie, Université de Mons-Hainaut, Avenue Maistriau, 19, B-7000 Mons, Bélgica), y a su vez traducidas a visual Basic por Francisco Pando.

2. GUÍA DE INSTALACIÓN Y USO

2.1. DESCARGA DEL PROGRAMA

Desde la página web: http://www.gbif.es/darwin_test/Darwin_Test.php se accede al punto de descarga de la aplicación:

1. Descargamos el fichero **Darwintest2-2k2.ZIP**.
2. Creamos un directorio donde instalar la aplicación (**recomendamos el directorio C:\Darwintest**) y descomprimos en dicho directorio el fichero ZIP que acabamos de descargar.
3. Para entrar en la aplicación, hacemos doble click sobre **Darwin_test.mde**
4. Aparece el menú principal:

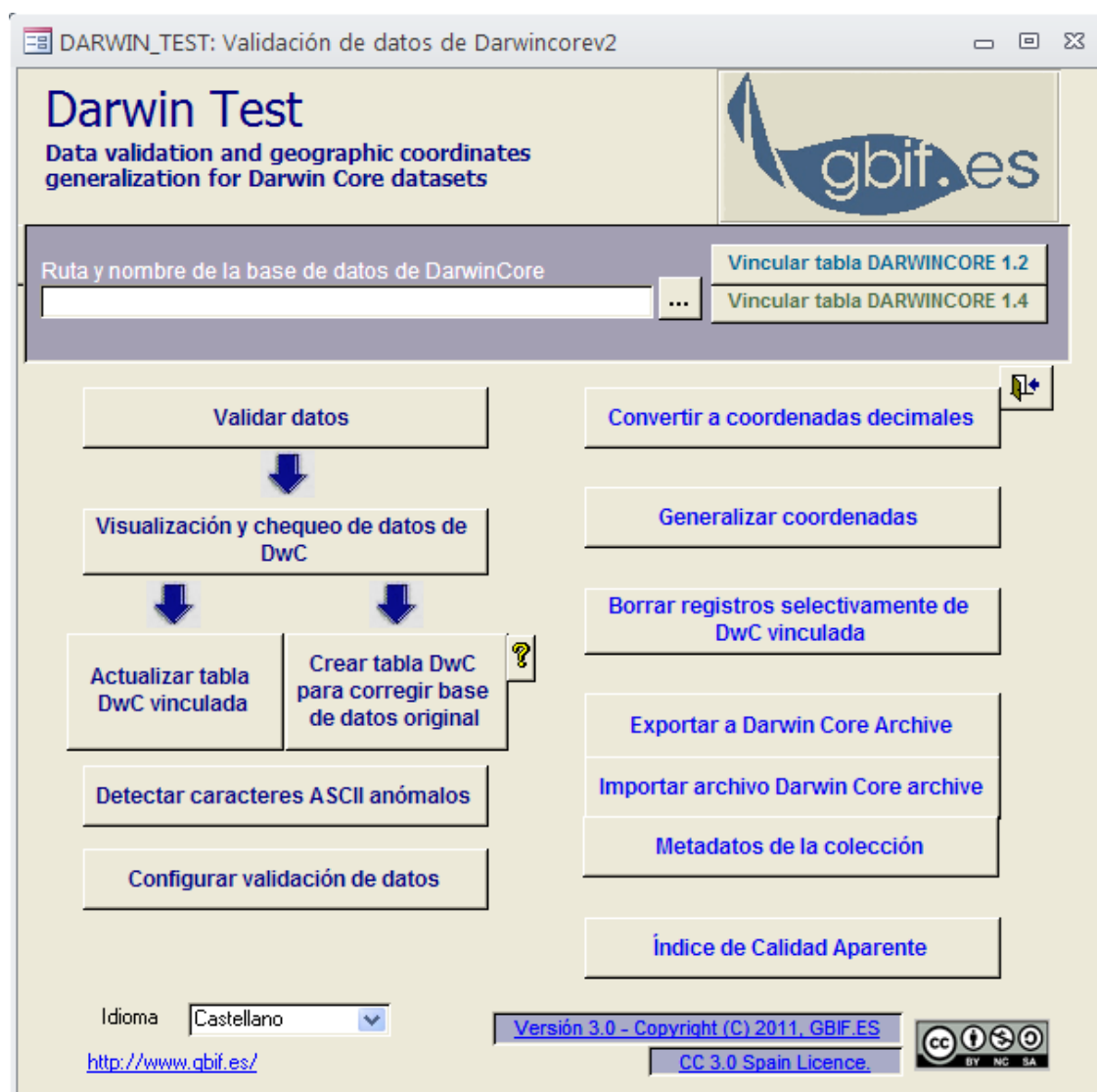


Figura 1.- Menú principal de DARWIN_TEST.

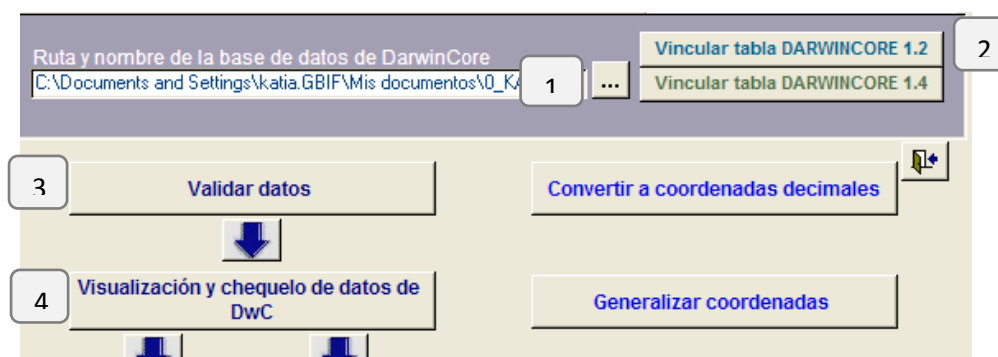
Puede configurarse el idioma desde el desplegable IDIOMA que se encuentra en la parte inferior del formulario.

2.2. VALIDAR LOS DATOS

Para validar nuestra tabla **DARWINCOREV2** basta con seguir estos 4 pasos:

- 1) Seleccionar con el botón de navegación [...] la base de datos en la que se encuentra nuestra tabla a validar, que debe llamarse **DARWINCOREV2**. En el caso de que nuestra tabla se llame diferente, pero tenga los nombres de los campos y su estructura igual al estándar **Darwincorev2**, sólo es necesario renombrar nuestra tabla y llamarla **DARWINCOREV2**.
- 2) Una vez seleccionada, pulsaremos en el botón **Vincular tabla DARWINCOREV2** que vinculará dicha tabla a nuestra base de datos de **Darwin_Test.mdb**.
- 3) Comenzaremos la validación pulsando en el botón **Validar datos**.
- 4) Tras finalizar la validación consultaremos 'los posibles' errores pulsando el botón **Visualización y chequeo de datos de DwC**.

Los cuatro pasos anteriores se corresponden con los botones numerados que se muestran en la Figura 2 y se aplican de igual modo si se quieren validar los datos de una tabla en formato **DARWINCOREV14**:

Figura 2.- Pasos para validar los datos de la tabla **DARWINCOREV2**.

Al pulsar el botón **Validar datos** saldrá el siguiente mensaje preguntándonos si deseamos actualizar el Índice de Calidad Aparente:

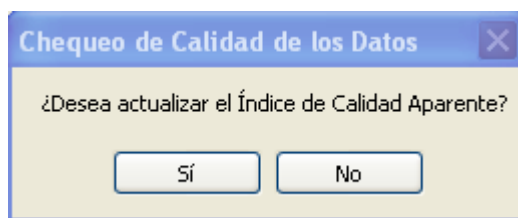


Figura 3.- Opción para actualizar el Índice de Calidad Aparente en cada validación.

Si la respuesta es **Sí**, entonces se guardan datos en la tabla **INDICE_CALIDAD** que pueden ser visualizados desde un formulario pulsando el botón **Índice de Calidad Aparente** del menú principal. Para saber más sobre este índice consulte el apartado [11. Índice de Calidad Aparente](#).

Al finalizar la validación de los datos pueden aparecer dos tipos de **mensajes**:

1. Mensaje con 'posibles' datos a corregir, que nos informa de las tablas que se han creado conteniendo los registros con POSIBLES datos erróneos:

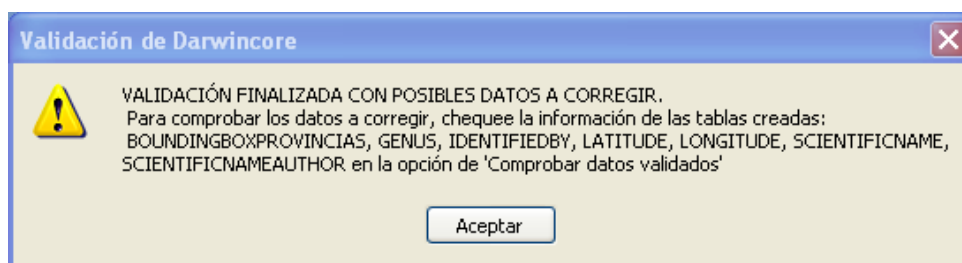


Figura 4.- Mensaje tras la validación de datos con 'posibles' errores a corregir.

Para ver los posibles errores debemos pulsar en el menú principal el botón **Visualización y chequeo de datos de DwC**. Así se abrirá un formulario que nos permite seleccionar el formulario con errores a revisar, tal como se explica en el siguiente apartado.

2. Mensaje comunicando la validación CORRECTA de los estándares, cuando no se han detectado errores.

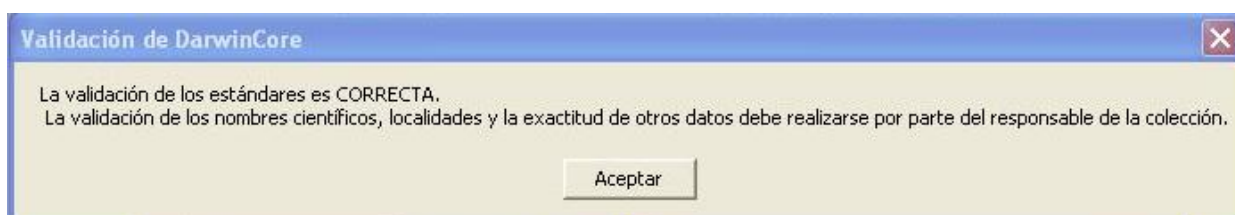


Figura 5.- Mensaje tras la validación de datos con estándares correctos.

IMPORTANTE. El hecho de que los estándares sean correctos no significa la ausencia de errores en los datos. Hay campos que no se validan debido a su complejidad para establecer unas reglas claras de validación. En otros casos, como en los nombres de las localidades, la exactitud de las coordenadas, el radio de incertidumbre, etc., es el usuario el que tiene que responder por la precisión y exactitud de los datos.

En cuanto a los nombres científicos, se pueden chequear contra una base de datos externa, como se explica en el apartado [4.3. Vincular tabla de Species2000](#).

Otros posibles errores que nos podrían aparecer (si hemos cambiado algo de la configuración) los podemos consultar en el capítulo [5. Control de los mensajes de error durante la validación](#).

Si la validación es correcta o el usuario la da por correcta, la tabla DARWINCORE está lista para ser subida al servidor correspondiente para su consulta y visualización en la web de GBIF.

IMPORTANTE. En caso de que se detecten errores, **es el usuario el que decide si los datos son correctos o no**. Es decir, no es necesario que salga el mensaje 2 de 'Validación Correcta', sino que una vez revisados los posibles errores, si el usuario decide que son correctos, puede dar por finalizado el proceso de chequeo y subir la tabla DARWINCOREV2 al servidor para su consulta y visualización en la Web.

Si la validación detecta la existencia de caracteres ASCII que se consideran anómalos dentro de un registro de base de datos, entonces nos aparecerá el siguiente mensaje:

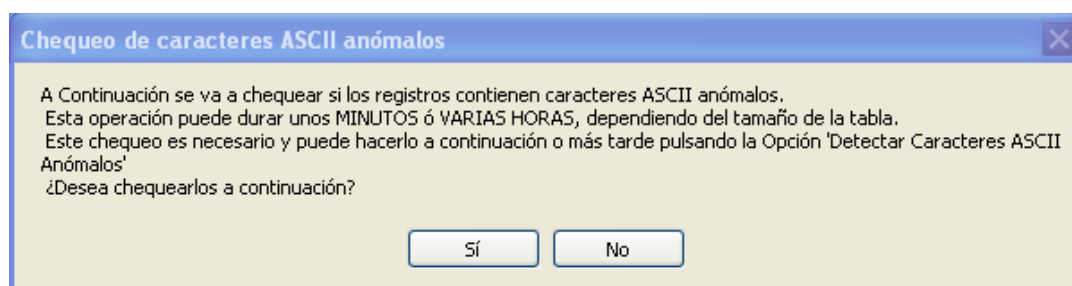


Figura 6.- Mensaje tras la detección de caracteres ASCII anómalos en algún campo del registro.

Estos caracteres pueden ir desde saltos de línea a tabuladores, salto de página, etc. Es importante que se eliminen de la tabla DARWINCORE ya que pueden interferir en el proceso de indexación de los datos en la web. El mensaje nos advierte de que, el proceso tarda desde varios minutos a

varias horas dependiendo del tamaño de la tabla y del tipo de ordenador que estemos utilizando. Como media, 40.000 registros tardan 40 minutos en procesarse y eliminar los caracteres ASCII anómalos encontrados. Si no disponemos de este tiempo, pulsaremos NO y revisaremos el resto de los datos chequeados. Cuando dispongamos de este tiempo, pulsaremos en la opción **Detectar caracteres ASCII anómalos**, antes de subir la base de datos al servidor web.

Para más información sobre este proceso de detección de caracteres ASCII anómalos, ver el capítulo [3. Detectar caracteres ASCII anómalos](#).

2.3. COMPROBAR LOS DATOS

Pulsando la opción **Visualización y chequeo de los datos de DwC** veremos todas las tablas temporales que se han creado y contienen los registros que tenemos que comprobar. No significa que todos los registros detectados tengan algún error, sólo es un aviso para que el usuario compruebe la veracidad y exactitud de los datos chequeados.

Para ver los datos detectados con posibles errores y corregirlos en caso de que lo sean, pulsaremos en cada uno de los campos que aparecen en el cuadro de la izquierda ‘Seleccionar la tabla para comprobar datos’ y pulsaremos en el botón **Chequear problemas encontrados**.

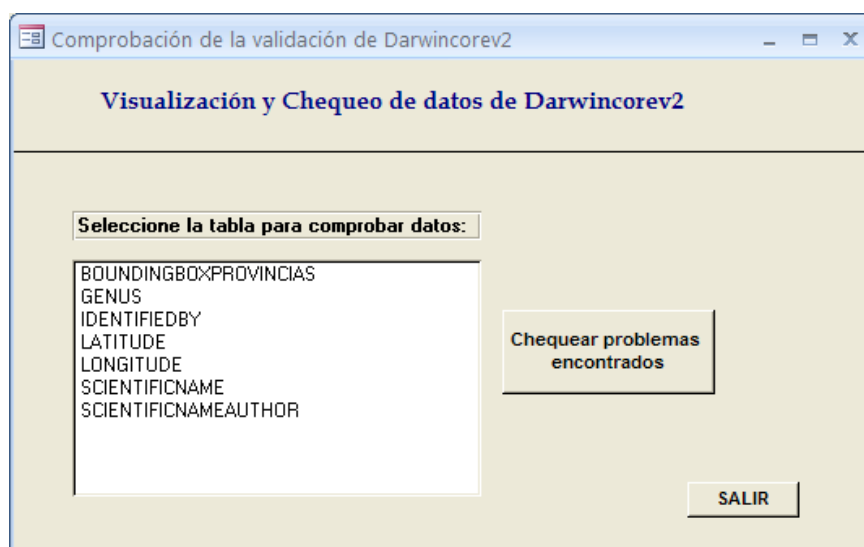


Figura 7.- Tablas temporales que contienen los datos a chequear.

Cada tabla seleccionada contiene los siguientes campos:

Ver Datos Chequeados de DARWINCOREV2

ScientificName:	lavatera cretica L.	Kingdom:	Plantae
Genus:	lavatera	Phylum:	
Species:	cretica	Class:	
Subspecies:		Order:	
ScientificNameAuthor:	L.	Family:	Malvaceae
IdentifiedBy:	J.A. Varela	TypeStatus:	
YearIdentified:		DayIdentified:	
InstitutionCode:	GBIF	CollectionCode:	Herbario_gbif
CatalogNumber:	1227-1	IndividualCount:	1
Collector:	J.A. Varela	BasisOfRecord:	S
CollectorNumber:		FieldNumber:	
YearCollected:	1978	MonthCollected:	5
JulianDay:		TimeOfDay:	
Country:	ES	ContinentOcean:	
StateProvince:	Co	County:	
Locality:	río Guadiato entre el puente de la ctra de Santa María de Trasierra y el cerro del Trigo		
Latitude:		CoordinatePrecision:	0
Longitude:		BoundingBox:	
MinimumElevation:	1	MaximumDepth:	0
MaximumElevation:	1	MinimumDepth:	0
Sex:		PreparationType:	Herbarium specimen
Notes:	esquistos Se espera que la primera letra del género sea en mayúsculas.		

PARA VALIDAR SALIR

Record: 1 of 1 No Filter Search

Figura 8.- Registro de la tabla seleccionada con todos sus campos.

En este formulario aparecen marcados con colores los siguientes campos:

- El **número de catálogo** (CatalogNumber) que identifica de forma única a cada registro está marcado en color naranja.
- El **campo validado** o detectado como posible error está recuadrado por un rectángulo de color rojo.
- En la parte inferior del formulario y remarcado en amarillo aparece un texto que explica **por qué ha sido seleccionado ese registro**. Esto NO significa que el dato sea erróneo, sólo quiere decir que ha sido chequeado con unas condiciones de validación y que no cumple la validación que se ha programado. El usuario es quien debe decidir si finalmente el dato es correcto o no.

Para desplazarnos entre los registros utilizaremos la barra de desplazamiento de Access situada en el pie del formulario

Record: 1 of 8 No Filter Search

2.4. CORREGIR LOS DATOS

A diferencia de las versiones anteriores de DARWIN_TEST, en las que las correcciones de los registros erróneos tenían que realizarse en la tabla Darwincorev2/Darwincore14 de origen, la nueva versión 2.0 permite corregir los datos desde el formulario de validación y traspasarlos a la base de datos Darwincorev2/Darwincore14 de una manera automática. Esta mejora agiliza tremendamente el proceso de depuración de la base de datos para su puesta en internet.

En el presente apartado se detalla el proceso de corrección de los datos y se describe la tabla intermedia que se crea antes del volcado de datos a cualquiera de las dos tablas anteriormente mencionadas.

La corrección puede realizarse directamente desde el mismo formulario de validación en el que se visualizan los registros problemáticos (Figura 8). Un aspecto a tener en cuenta, derivado del diseño de la aplicación, es que **únicamente se guardarán los cambios en aquel o aquellos campos que en el formulario estén marcados en azul con un recuadro verde**, y que coincide con el campo o campos que se están validando.

Ejemplo. En la figura 8 podemos observar un caso en el que ha sido chequeado el campo GENUS y aparece marcado en azul con un recuadro verde en el formulario, además, el campo SCIENTIFICNAME tiene el mismo error tipográfico, y es que el género comienza con minúscula. En este caso la corrección habría que realizarla únicamente en el campo resaltado ya que cualquier rectificación en alguno de los otros campos no quedaría registrada en la tabla final.

Una vez que se han realizado las correcciones pertinentes en un registro dado, hay que pulsar el botón **Para validar**, de esta manera las correcciones se van almacenando en una tabla intermedia llamada **CAMPOSTABLASVALIDADOS**. Ésta recoge la información sobre los registros y campos que sufren correcciones: el nombre de la tabla de origen resultado de la validación está especificada en el campo **NOMBRE_TABLA**. El registro en el que se van a realizar los cambios está especificado en la columna **KEY**. Las columnas encabezadas con una **C**, nos informan de si para esa tabla se van a corregir uno, dos o tres campos. Las columnas encabezadas con una **V**, sólo se rellenan en algunos casos con los nuevos valores para ese registro.

NOMBRE_TABLA	PARA_VALIDA	C	V1	C	V2	C3	V3	Key	Fecha_Actualizaci
a_SCIENTIFICNAME	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input type="checkbox"/>		<input type="checkbox"/>		10000-1	25/01/2010 9:28:31
a_GENUS	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Lavatera	<input type="checkbox"/>		<input type="checkbox"/>		1227-1	25/01/2010 8:22:45
	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>		<input type="checkbox"/>			

Figura 9.- Tabla CAMPOSTABLASVALIDADOS en la que se almacena la información de los registros y sus campos que han sufrido correcciones.

2.5. TRASPASAR LOS DATOS CORREGIDOS A LA TABLA DARWINCORE

Los cambios anteriores permanecerán registrados en la tabla intermedia

CAMPOSTABLASVALIDADOS y no serán efectivos hasta que se pulse alguna de estas dos opciones del menú principal:

- **Actualizar tabla DwC vinculada:** realiza los cambios de manera permanente en la tabla Darwincorev2/14 vinculada y por tanto afecta también a la tabla original.
- **Crear tabla DwC para corregir la base de datos original:** crea una tabla gemela de Darwincorev2/14 que contiene las correcciones realizadas por el usuario además de otros campos informativos. Dicha tabla se llama Darwincorev2Validada/Darwincorev14Validada.

IMPORTANTE. Las correcciones realizadas a través de cualquiera de estos dos procesos únicamente guarda los cambios en la tabla Darwincorev2/14 o Darwincorev2Validada/Darwincorev14Validada con lo que los datos aparecerán de una manera correcta en la web, pero para asegurar la calidad de nuestros datos dichos cambios deberían realizarse también en la base de datos original.

IMPORTANTE.

- ✓ En la validación de las coordenadas de Latitude y Longitude para comprobar si caen dentro del Bounding box, hay que tener en cuenta que los valores de estos rectángulos o Boundingbox que encuadran a cada país, no están incluidos los límites de las aguas territoriales de cada país para los organismos marinos.
- ✓ En otros casos, es posible que no estén incluidas las islas, por ejemplo, en el Boundingbox de Estados Unidos no está incluido Hawai. Si se desea, se pueden ampliar los límites del Boundingbox de cada país modificando los valores de la tabla ISO-3166 (campos bBoxNorte, bBoxSur, bBoxEste, bBoxOeste).

3. DETECTAR CARACTERES ASCII ANÓMALOS

En ocasiones los datos de la tabla **DARWINCORE** tienen caracteres extraños que pueden provenir de distintas fuentes: importación de datos desde documentos de textos con formato, desde hojas de cálculo, uso de teclas especiales durante la captura de datos, etc. Estos caracteres pueden variar desde saltos de líneas a tabuladores verticales, tabuladores horizontales, etc., y pueden dar lugar a problemas en los procesos de reindexación de los datos en los procesos realizados en el portal de GBIF internacional (<http://de.mirror.gbif.org/portal/welcome.htm>).

Para evitar que estos caracteres dañen los procesos de reindexación, conviene detectarlos y reemplazarlos por caracteres válidos. Esto se realiza automáticamente pulsando en la opción **Detectar caracteres ASCII anómalos**. Este proceso puede durar varias minutos o varias horas, dependiendo del tamaño de la tabla Darwincore (**aproximadamente 1.000 registros se chequean en 1 minuto**, por lo que por ejemplo, una tabla de 30.000 registros tardaría 30 minutos).

Por ello, se recomienda realizarlo después de la validación de datos o junto con ella, ya que durante la misma nos pregunta si queremos realizar ambos procesos, pero teniendo en cuenta que puede estar bastantes minutos e **incluso HORAS** realizando esta detección de caracteres. Cuando acaba el proceso de detección de caracteres anómalos sale el siguiente mensaje.

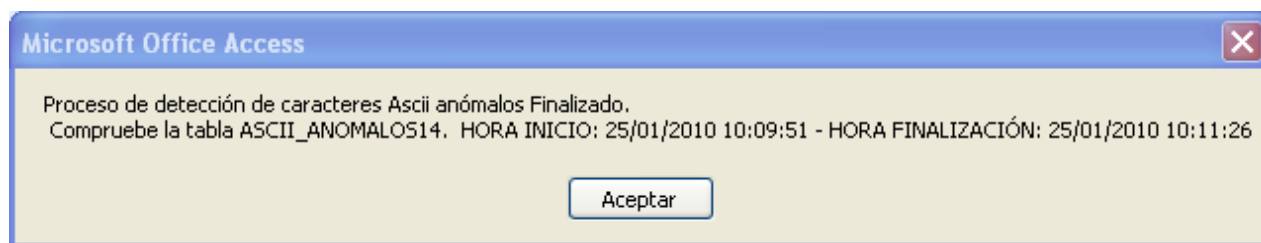


Figura 10.- Aviso de finalización de la detección de caracteres ASCII anómalos.

Si se detectan caracteres ASCII extraños serán reemplazados automáticamente en la tabla **DARWINCORE**, pero tendremos que realizar el mismo cambio en los datos originales de nuestra colección, ya que de lo contrario, no eliminaremos el origen de dichos caracteres ASCII anómalos. Para saber qué caracteres ASCII extraños o anómalos tienen la tabla **DARWINCORE**, al finalizar el proceso de detección abriremos la tabla **ASCII_ANOMALOS**, que contendrá registros en los que se han detectado los caracteres detectados antes de reemplazarlos en **DARWINCORE**.

Ord	Family	Genus	SpecificEpithet
	Tricholomataceae	Tricholoma	gausapatum
	Arcyriaceae	Arcyria	cinerea
	Didymiaceae	Mucilago	crustacea
	Echinosteliaceae Echinosteliaceae	Echinostelium	cribrarioides
	Dictydiaethaliaceae Dictydiaethaliaceae	Dictydiaethalium	plumbeum
	Dictydiaethaliaceae Dictydiaethaliaceae	Dictydiaethalium	ferrugineum
	Stemonitaceae Stemonitaceae Stemonitaceae	Comatricha	nigra

Figura 11.- Registros con caracteres ASCII inválidos para su visualización en la web.

La sustitución de estos caracteres inválidos por otros válidos se realiza automáticamente en función de los valores que aparecen en la tabla **SUSTITUIR_ASCII**, que el usuario puede cambiar según sus necesidades:



Caracter_Ascii	Sustitucion	Descripcion
9	.	Retroceso
10	.	Tabulador horizontal
11	-	Tabulador vertical
12	.	Salto de linea
13	.	Retorno de carro
16	.	Escape
27	.	Escape
126	.	

Figura 12.- Tabla SUSTITUIR_ASCII. Muestra el tipo de carácter ASCII extraño (descripción) y el valor por el que será sustituido (Sustitución).

4. CONFIGURAR Y PERSONALIZAR LA APLICACIÓN

4.1. CONFIGURAR LA TABLA DE VALIDACIONES

Las validaciones que se realizan en el botón **Visualización y chequeo de los datos de DwC** son consultas de Access que se pueden ver cuando pulsamos la tecla **F11** de nuestro teclado y seleccionamos la sección de 'Consultas' dentro del menú de Access:

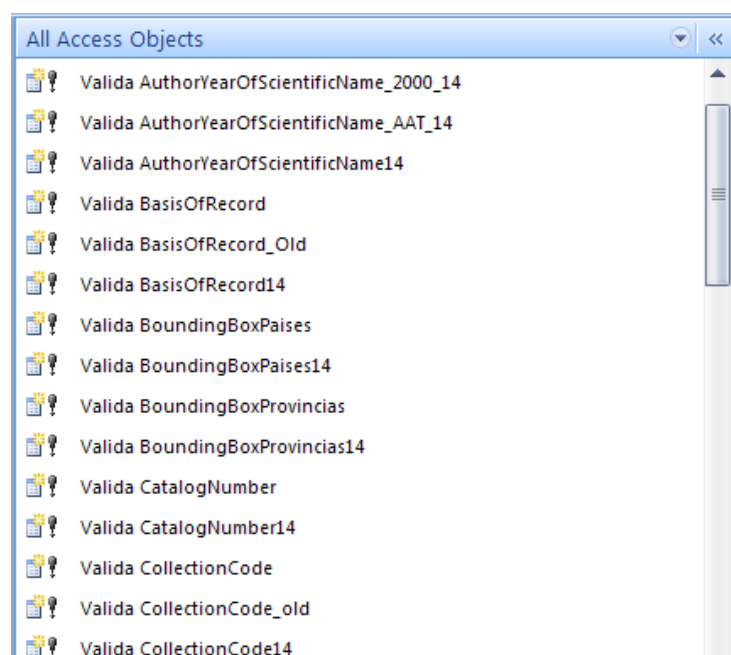


Figura 13.- Consultas o 'Queries' de validación de datos que vienen predeterminadas en la aplicación.

Estas consultas están activadas por defecto y se ejecutan al pulsar el botón **Visualización y chequeo de los datos de DwC**. Una vez que se ejecutan, cada una de ellas crea como resultado una tabla que empieza por **A_** seguido por el nombre del campo que valida. Dichas tablas contienen los registros que se han encontrado con algún posible error en el campo en cuestión y que son los mismos que se visualizan en el formulario descrito en el apartado [2.3. Comprobar los datos](#).

Todas estas consultas vienen activadas por defecto pero si se desea, pueden personalizarse según las necesidades de cada usuario. Para ir al menú de configuración desde donde se puede realizar esta operación, hay que pulsar el botón **Configurar validación de datos** del menú principal de Darwin Test que dará paso al siguiente formulario:



Opciones de configuración de Darwin_Test

CONFIGURACIÓN DE VALIDACIÓN DE DATOS

Elija la configuración que desea realizar:

Ruta y nombre de la base de datos Species 2000 C:\Darwin_test\Species2000_Scientific_names.mdb	...	Vincular tabla SCIENTIFIC_NAMES - Species 2000
Ruta y nombre de la base de datos de AAT C:\Darwin_test\AATs del SIB Colombia.mdb	...	Vincular tabla AAT - Archivos Autoridad Taxonómica

Configurar la tabla de Validaciones para Darwincore1.2

Configurar la tabla de Validaciones para Darwincore1.4

SALIR

Figura 14.- Opciones de configuración de Darwin Test.

En la parte inferior hay dos botones que permiten el acceso a los formularios de configuración de las validaciones. Pulsaremos uno u otro en función de si queremos configurar las validaciones para la tabla Darwincorev2 o Darwincore1.4. En ambos casos el formulario presenta los campos que se explican a continuación.

CAMPO A VALIDAR	CONSULTA QUE LANZA	ACTIVO	VALIDACIONES QUE REALIZA
CatalogNumber	Valida CatalogNumber	<input checked="" type="checkbox"/>	Su valor no puede ser nulo ni estar repetido dentro de una misma colección.
Class		<input type="checkbox"/>	
CollectionCode	Valida CollectionCode	<input checked="" type="checkbox"/>	Su valor no puede ser nulo y es recomendable tener una única colección por cada tabla de Darwincore.
Collector	Valida Collector	<input checked="" type="checkbox"/>	Se espera que el nombre del colector esté escrito con la primera letra en mayúsculas.
CollectorNumber		<input type="checkbox"/>	
ContinentOcean	Valida ContinentOcean	<input checked="" type="checkbox"/>	El nombre del continente debe existir en la tabla Continentes.
CoordinatePrecision	Valida CoordinatePrecision	<input checked="" type="checkbox"/>	Se espera un valor numérico entre 0 y 80.000

Record: 5 of 60 No Filter Search

Figura 14.- Formulario de configuración de validaciones.

- **Campo a Validar:** el nombre del campo de la tabla Darwincorev2 o Darwincore 1.4 que la consulta va a comprobar.
- **Consulta que lanza:** nombre de la consulta de Access incluida en la aplicación que va a realizar la validación y que está incluida en la sección de consultas de Access.
- **Activo:** esta casilla indica si la consulta en cuestión se va a ejecutar (activada) o no (desactivada)
- **Validaciones que realiza:** describe la validación que realiza esa consulta, para nuestra información y para que salga como aviso en los registros que no cumplan dicha validación.

El usuario puede limitar las consultas que lanza la aplicación activando o desactivando el campo 'Activo' del formulario anterior para cada caso. Igualmente puede cambiar la descripción de las consultas ya existentes e incluso añadir consultas propias que previamente habrán sido diseñadas y guardadas en la sección de consultas de Access.

EJEMPLO DE CREACIÓN DE NUEVA CONSULTA DE VALIDACIÓN

Ejemplo: queremos validar el campo **INDIVIDUALCOUNT**, que corresponde al número de ejemplares de ese registro. La validación va a consistir en que nos dé un mensaje de error si el número de ejemplares es mayor de

15, por ejemplo. Para hacer esto, pulsamos en el menú de Access lo siguiente:

Consultas → Nueva → Vista Diseño → Mostrar tabla

y añadimos la tabla **DARWINCOREV2**. De ésta, seleccionamos todos los registros pulsando en Campo (Field) y seleccionamos **DARWINCOREV2.*** En otra casilla de la consulta seleccionamos el campo

INDIVIDUALCOUNT y desmarcamos la casilla de **Mostrar** y ponemos el **Criterio:** > 15 (Ver Figura 15).

A continuación, sobre la parte gris/azul de la consulta pulsamos el botón derecho del ratón y nos muestra el menú donde podemos elegir el Tipo de consulta (“Query Type”) y elegimos ‘Consulta de Datos anexados’ (Append Query) y lo anexamos a una tabla que llamaremos

A_INDIVIDUALCOUNT:

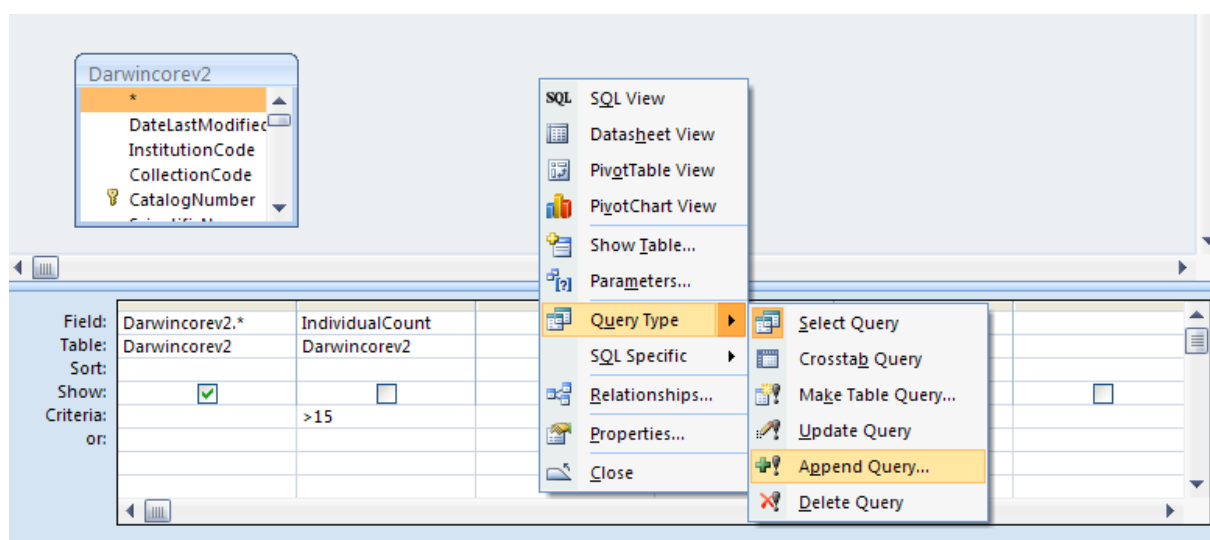


Figura 15.- Creación de una consulta de validación del tipo Anexión de Datos.

Guardamos la consulta como ‘**Valida IndividualCount**’ y cerramos el diseño.

Una vez que tenemos creada la consulta hay que agregarla al formulario de 'Configuración de validaciones'. Para ello pulsamos el botón **Configurar validación de datos**, y el botón **Configurar la tabla de Validaciones para Darwincorwv2/1.4** según cual sea el caso.

En el siguiente formulario que nos aparece nos movemos con la barra de desplazamiento vertical y buscamos el campo a validar INDIVIDUALCOUNT. En este campo deben rellenarse los siguientes valores:

CAMPO A VALIDAR	CONSULTA QUE LANZA	ACTIVO	VALIDACIONES QUE REALIZA
IndividualCount	Valida IndividualCount	<input checked="" type="checkbox"/>	Espera que el número de ejemplares del espécimen catalogado sea menor que 15.

Figura 16.- Configurar la validación de un campo que carecía de validación.

- **Consulta que lanza:** escribimos nombre de la consulta que hemos creado en Access y que hemos llamado **Valida IndividualCount**.
- **Activo:** activamos esta casilla para que la consulta se ejecute en el momento de la validación de los datos.
- **Validaciones que realiza:** escribimos lo que hace esta validación, para nuestra información y para que salga como aviso en los registros que no cumplan dicha validación.

La siguiente vez que realicemos la validación de los datos, se ejecutará la consulta **Valida IndividualCount** y si hay registros que exceden el número de Individuos que hemos puesto, entonces se creará la tabla A_INDIVIDUALCOUNT y podremos consultar dichos registros en la opción **Visualización y chequeo de datos de DwC**.

IMPORTANTE.

- ✓ También se puede realizar cada validación por separado, es decir, si pulsamos la tecla F11 y pulsamos en Consultas de Access, podemos elegir la consulta deseada, por ejemplo, '**Valida IndividualCount**' y al hacer doble-click sobre ella nos valida únicamente el campo INDIVIDUALCOUNT. En esta validación nos creará la tabla **A_INDIVIDUALCOUNT** siempre, aunque el número de errores sea cero, en cuyo caso se creará la tabla con cero registros.
- ✓ También se pueden modificar las consultas de validación existentes e incluir nuevos criterios de validación según las necesidades de cada usuario.

4.2. ELEGIR EL GRADO DE GENERALIZACIÓN DE LAS COORDENADAS


En ocasiones, o en el caso de ciertos registros o especies, puede interesar presentar las coordenadas con cierto grado de generalización o eliminarlas para que estos datos no sean publicados en la web. Existe un formulario específico para ello en DARWIN TEST, al que podemos acceder desde el botón **Generalizar coordenadas**. Dentro de éste, existen varias opciones para aplicar distinto grado de generalización o para eliminar las coordenadas, previa elección de los registros a aplicar dicha función. Su funcionamiento se explica más adelante en el apartado [7. Difuminar o Eliminar de coordenadas de Darwincorev2](#).

4.3. VINCULAR LA TABLA DE SPECIES2000 O ARCHIVOS DE AUTORIDAD TAXONÓMICA

Darwin Test ofrece la posibilidad de comprobar la validez de los nombres científicos contra la tabla de nombres de **CATALOGUE OF LIFE** (Species 2000) o la tabla de los **Archivos de Autoridad Taxonómica (AAT)**. Estas tablas pueden ser descargadas en formato Access desde los vínculos que pueden encontrarse en la zona de descarga de Darwin Test:

http://www.gbif.es/darwin_test/Darwindownload.php.

Si disponemos de estas tablas en formato Access (.mdb) podemos vincular la tabla de nombres científicos y realizar las validaciones de género, especie y subespecie contra dichas base de datos. Para hacerlo, dentro de la opción **Configurar validación de datos**, primero buscamos el

directorio donde tenemos la base de datos CATALOGUE OF LIFE que hemos descargado previamente, y navegamos con el botón  para encontrarla en el directorio en el que la hemos almacenado. Una vez seleccionada, pulsamos el botón **Vincular tabla SCIENTIFIC-NAMES de Species 2000**.

El proceso para vincular la tabla AAT.

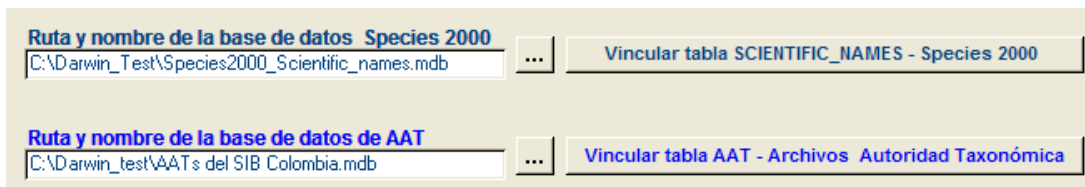
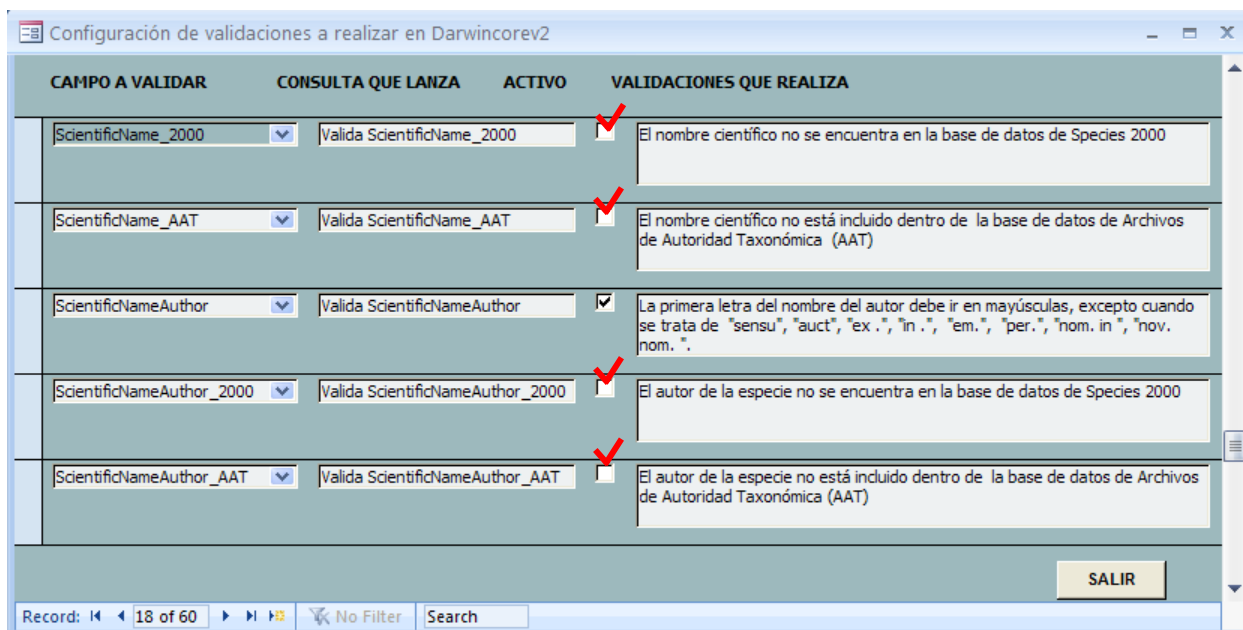


Figura 17.- Formulario de vinculación de bases de datos externas Catalogue of Life y ATT

Para realizar comprobación de nuestros nombres científicos con los de estas tablas existen varias consultas de validación para las tablas Darwincorev2 y Darwincore1.4. Dichas consultas vienen desactivadas por defecto, y en caso de que el usuario desee que se ejecuten automáticamente al pulsar el botón de **Visualización y chequeo de datos de DwC**, tiene que activarlas en el formulario de 'Configuración de validaciones' tal y como se explicó en el apartado [4.1. Configurar la tabla de Validaciones](#) y como puede verse en la siguiente figura:



CAMPO A VALIDAR	CONSULTA QUE LANZA	ACTIVO	VALIDACIONES QUE REALIZA
ScientificName_2000	Valida ScientificName_2000	<input checked="" type="checkbox"/>	El nombre científico no se encuentra en la base de datos de Species 2000
ScientificName_AAT	Valida ScientificName_AAT	<input checked="" type="checkbox"/>	El nombre científico no está incluido dentro de la base de datos de Archivos de Autoridad Taxonómica (AAT)
ScientificNameAuthor	Valida ScientificNameAuthor	<input checked="" type="checkbox"/>	La primera letra del nombre del autor debe ir en mayúsculas, excepto cuando se trata de "sensu", "auct", "ex .", "in .", "em.", "per.", "nom. in ", "nov. nom. "
ScientificNameAuthor_2000	Valida ScientificNameAuthor_2000	<input checked="" type="checkbox"/>	El autor de la especie no se encuentra en la base de datos de Species 2000
ScientificNameAuthor_AAT	Valida ScientificNameAuthor_AAT	<input checked="" type="checkbox"/>	El autor de la especie no está incluido dentro de la base de datos de Archivos de Autoridad Taxonómica (AAT)

Figura 18.- Activar las validaciones de Species2000

En DARWIN_TEST, en la tabla **VALIDACIONES**, el nombre de 'CAMPO A VALIDAR' se llama SCIENTIFICNAME_2000, GENUS_2000, SPECIE_2000 y SUBESPECIES_2000 para distinguirlo de las otras validaciones de GÉNERO, ESPECIE y SUBESPECIE, que validan ciertas normas respecto del código de nomenclatura. Ocurre lo mismo para el caso de los campos de la tabla AAT: SCIENTIFICNAME_AAT, GENUS_AAT.

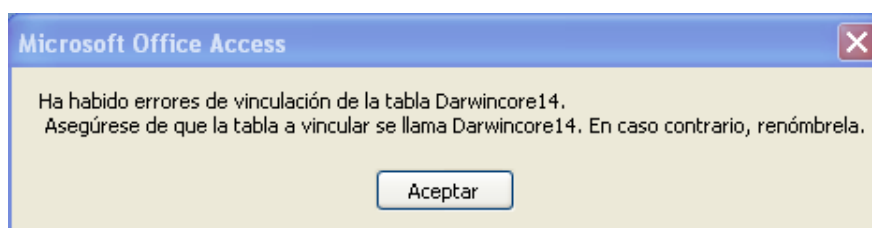
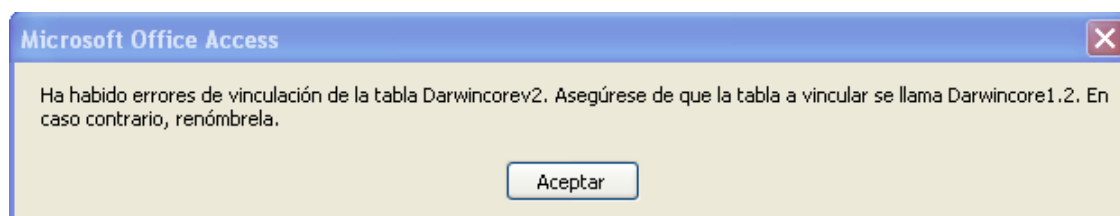
IMPORTANTE. Si en vez de la tabla de nombres de Species2000 & ATIT tenemos otra tabla de nombres de referencia, podremos utilizarla de la misma manera, vinculándola igual que acabamos de ver pero con la única condición de que la tabla se llame **SCIENTIFIC_NAMES** y los campos a validar en dicha tabla se llamen GENUS, SPECIES y SUBSPECIES.

5. CONTROL DE MENSAJES DE ERROR DURANTE LA VALIDACIÓN

Los mensajes de error durante la validación de datos pueden producirse debido a que hayamos realizado algún cambio en la configuración y éste cambio no se haya realizado bien. Hay dos posibles casos de error que nos detecta la validación.

ERROR 1. Mensaje de error por nombre incorrecto en la tabla Darwincore a vincular.

Si el nombre de la tabla Darwincore que desea vincular es erróneo le saldrá alguno de los siguientes mensajes:



Figuras 19 y 20.- Mensajes de error durante la vinculación de la tabla Darwincore.

ERROR 2. Mensaje de error por especificar la consulta que realiza la validación, pero esta consulta no ha sido creada todavía.

Si durante la configuración de la tabla de VALIDACIONES (botón **Configurar la tabla de validaciones**) hemos activado un campo a validar y le hemos dicho el nombre de la consulta que valida dicho campo, pero dicha consulta no existe o su nombre está mal escrito nos dará el siguiente mensaje:

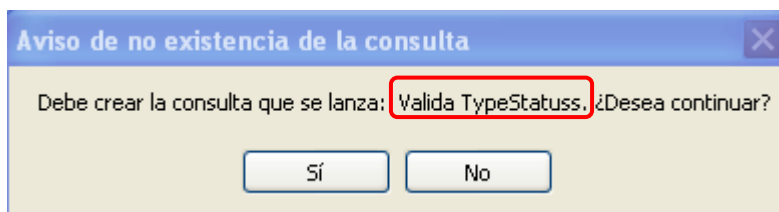


Figura 21.- Mensaje de error durante la validación

En este ejemplo, tendremos que ir a la sección de consultas de Access y comprobar que realmente existe dicha consulta o qué su nombre está bien escrito en la tabla de validaciones. Si no existe habrá que crearla como se ha explicado en apartados anteriores.

ERROR 3. Mensaje de error por no especificar el nombre de la consulta que realiza la validación.

Si durante la configuración de la tabla de VALIDACIONES (botón **Configurar la tabla de validaciones**) hemos Activado un campo a validar pero no le hemos dicho el nombre de la consulta que tiene que validar dicho campo, entonces no dará el mensaje:

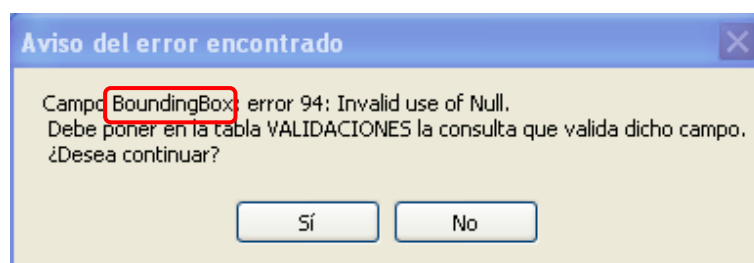


Figura 22.- Mensaje de error durante la validación

En este ejemplo, tendremos que ir a la configuración de la tabla de validaciones y buscar el campo BOUNDINGBOX, y en la casilla 'Consulta que lanza' especificar el nombre de la consulta de

validación de dicho campo, que previamente habremos creado en Access como se explicó anteriormente.

6. CONVERTIR COORDENADAS A COORDENADAS DECIMALES

Desde el botón **Convertir a coordenadas decimales** podemos transformar coordenadas en distintos sistemas a coordenadas geográficas en grados decimales, que es el formato requerido en Darwincorev2 para que los registros georreferenciados se visualicen en los mapas de distribución de la web del GBIF. Si los registros de Darwincorev2/Darwincore14 no tienen las coordenadas en este formato dentro de los campos LATITUDE y LONGITUDE, entonces se debe utilizar esta opción para transformar las coordenadas originales.

Para ello, las coordenadas de origen deben estar incluidas en un campo con el nombre MISCOORDENADAS, de tipo Texto en la tabla Darwin Core. Si el campo donde tenemos las coordenadas se llama de otra forma, lo renombraremos. En él se pueden especificar tanto coordenadas **geográficas** como **UTM's** tanto del tipo **MGRS** como especificando el ***Easting*** y ***Northing***.

- En el caso de UTM del tipo **MGRS** (Military Grid Reference System) se introducen los caracteres sin espacios entre ellos. Ejemplo: 30TUV4050.
- En el caso de UTM con ***Easting*** y ***Northing***, se introducen los datos dejando un espacio entre la Zona y el *Easting*, y entre el *Easting* y el *Northing*. Ejemplo: 30T 440 4650.
- En el caso de **coordenadas geográficas en grados sexagesimales** el formato es Grados°minutos'segundos"[NS] Grados°minutos'segundos"[EW], dejando un espacio en blanco entre la latitud y la longitud. Ejemplo: 30°50'15"N 2°30'10"W.
- En el caso de **coordenadas geográficas en grados decimales** el formato debe ser [-]latitud [-]longitud, dejando un espacio entre ellos e indicando en cada caso su correspondiente signo. Ejemplo: 40.3388 -2.0220.

A continuación pulsaremos en el botón de Darwin Test **Convertir a coordenadas decimales**, y nos saldrá el siguiente mensaje:

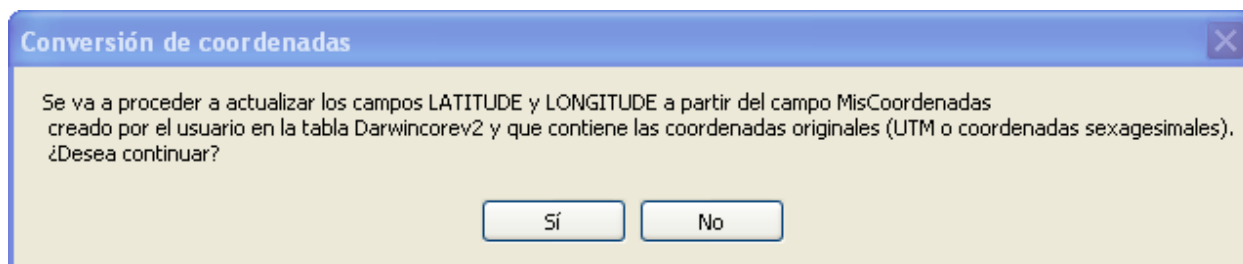


Figura 23.- Mensaje de aviso para la conversión de coordenadas.

Si pulsamos **SÍ** convertirá las coordenadas del campo MISCOORDENADAS (que hemos creado o renombrado previamente en la tabla **DARWINCOREV2**) a coordenadas decimales en los campos LATITUDE y LONGITUDE. A continuación el programa nos hará la siguiente pregunta:

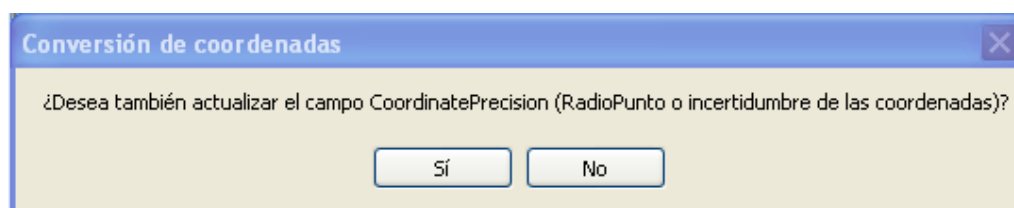


Figura 24.- Mensaje de aviso para la conversión de coordenadas.

Pulsaremos **SÍ** cuando el campo COORDINATEPRECISION (equivalente al campo RadioPunto o radio de incertidumbre de las coordenadas) no contenga la información correcta o esté vacío y será, por tanto, necesario actualizarla. Diremos que **NO** cuando la información del campo COORDINATEPRECISION la hayamos introducido previamente y queramos que permanezca esta información y no la calculada por el programa.

Debemos comprobar que las coordenadas de los campos de LATITUDE y LONGITUDE que acabamos de convertir tengan **decimales**, ya que de lo contrario la visualización de las coordenadas en el mapa sería muy poco precisa y los puntos pueden quedar muy alejados de las coordenadas reales y/o pueden caer en el mar (cuando corresponden a organismos terrestres) o viceversa.

IMPORTANTE 1. Para puntos del ‘Viejo mundo’, la conversión utiliza el elipsiode WGS84, lo que da un error de unos 200 m con respecto a las coordenadas calculadas sobre los mapas de IGN o del SGE, que utilizan el Datum ‘European 50’. Para puntos de las Américas se utiliza el elipsiode ‘Clarke 1866’, que es el que utilizan la mayoría de los mapas de esos pagos. (Nota de Francisco Pando: Noviembre 2007)

IMPORTANTE 2. Para que las coordenadas en grados decimales salgan con decimales (Latitude 42,36 por ejemplo) es necesario que el separador decimal de Windows sea **el punto** y no la coma. Para cambiarlo, pulsar en el botón **Inicio** de Windows → **Configuración** → **Panel de Control** → **Configuración regional** → **Números** → **Separador decimal** → aquí pondremos el **punto** en vez de la coma.

7. DIFUMINAR O ELIMINAR COORDENADAS DE DARWINCORE

La opción de difuminar las coordenadas se puede realizar si no se desea informar de la verdadera localización del espécimen.

Antes de realizar esta acción es conveniente **realizar una copia de seguridad** de la tabla **DARWINCOREV2** original tanto si se van a difuminar como si se van a eliminar las coordenadas.

IMPORTANTE. NO debemos realizar **nunca** una copia de seguridad de una tabla **vinculada**, porque no es una verdadera copia, sino que se copia el acceso directo nada más y en realidad está apuntando a la misma tabla. Por ello, se debe hacer siempre la copia de seguridad de la tabla **DARWINCOREV2** en la base de datos en donde **reside realmente dicha tabla**.

La operación se realiza desde el botón **Generalizar coordenadas**, que abre un formulario desde el que tenemos que establecer las condiciones para seleccionar los registros cuyas coordenadas deseamos difuminar.



Figura 25.- Selección de los registros cuyas coordenadas deseamos difuminar o eliminar.

Para seleccionar los registros, debemos elegir las condiciones de la selección. Si queremos concatenar varias condiciones, entonces utilizaremos el campo UNION para enlazar con la siguiente condición. Pulsando el botón **Seleccionar registros** tendremos los registros a tratar. Podemos ver los registros seleccionados en forma de consulta o formulario pulsando a los botones **Ver consulta** y **Ver formulario** respectivamente. Si hemos utilizado esta opción en otra ocasión, entonces deberemos pulsar previamente **Deseleccionarlos** ya que es un filtro o selección acumulativa, de manera que una selección se va realizando sobre la selección anterior.

Una vez que tenemos los registros deseados seleccionados, pulsaremos en la opción que deseemos realizar:

- **Generalizar coordenadas.** Podemos aplicar distintito grado de generalización a las coordenadas: **0.1**, **0.01** o **0.001**, que equivaldría a una precisión aproximada de 10 Km, 1 Km y 100 m respectivamente. Al pulsar cualquiera de estas opciones se calculara de forma automática el valor del campo COORDINATEPRECISION o radio de incertidumbre de la coordenada.

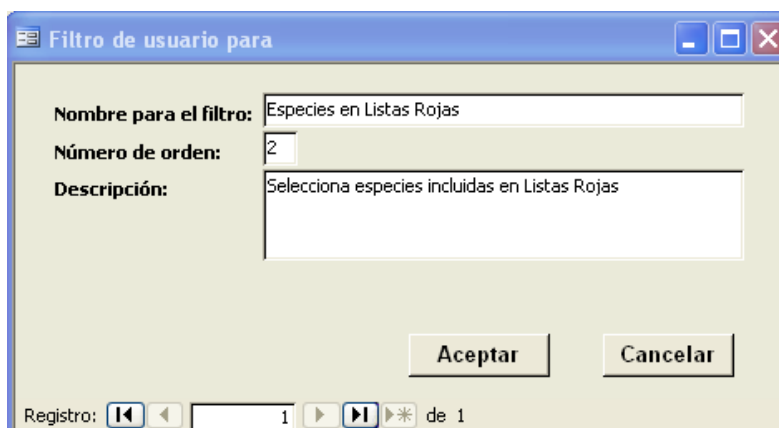
Es conveniente que tras esta generalización se realice de nuevo la validación de los datos, ya que las coordenadas pueden caer fuera del Boundingbox o rectángulo que encuadra al país o la provincia, o que incluso sobrepasen los valores máximos de latitud (+/- 90°) o longitud (+/- 180°). También se puede realizar la validación independiente de cada coordenada, pulsando la tecla F11, pulsando en Consultas de Access, seleccionando las consultas (**Valida Latitude** y **Valida Longitude**) y haciendo doble-click sobre cada una de ellas.

- **Eliminar coordenadas y radiopunto.** Borra el valor de las coordenadas de los campos Latitude y Longitude de la tabla **DARWINCOREV2**.

IMPORTANTE 2. Si se van a difuminar las coordenadas, **será necesario informar siempre de este hecho a los usuarios que vayan a utilizar o consultar estos datos.** Es decir, el hecho de la generalización de las coordenadas **debe ser claramente expuesto en los metadatos de la colección**, para que las personas que vayan a consultar los datos estén informadas de este hecho y no dé lugar a una falta de veracidad ni inexactitud de los datos.

ALMACENAMIENTO DE FILTROS

También existe la posibilidad de almacenar los filtros si es que se desean utilizar en otras ocasiones o tienen condiciones de selección de registros complejas. Para ello, una vez establecidas las condiciones pulsamos el botón **Guardar filtro actual** que abre el siguiente formulario:



El formulario, titulado "Filtro de usuario para", contiene los siguientes campos:

- Nombre para el filtro:** Especies en Listas Rojas
- Número de orden:** 2
- Descripción:** Selecciona especies incluidas en Listas Rojas

En la parte inferior del formulario hay dos botones: "Aceptar" y "Cancelar".

En la parte inferior del formulario, se muestra un control de registro: "Registro: 1 de 1" con botones de navegación.

Figura 26.- Almacenamiento de nuevo filtro.

Aquí, damos un nombre y una descripción del filtro y tras pulsar **Aceptar**, el filtro queda almacenado y disponible para otra ocasión.

Para **gestionar y configurar los filtros** almacenados pulsamos el botón **Gestión de filtros**, donde podemos editar el nombre, el tipo de acción y la descripción del filtro. Así mismo podemos activar o desactivar el filtro para que nos quede disponible en el formulario de selección de filtros.

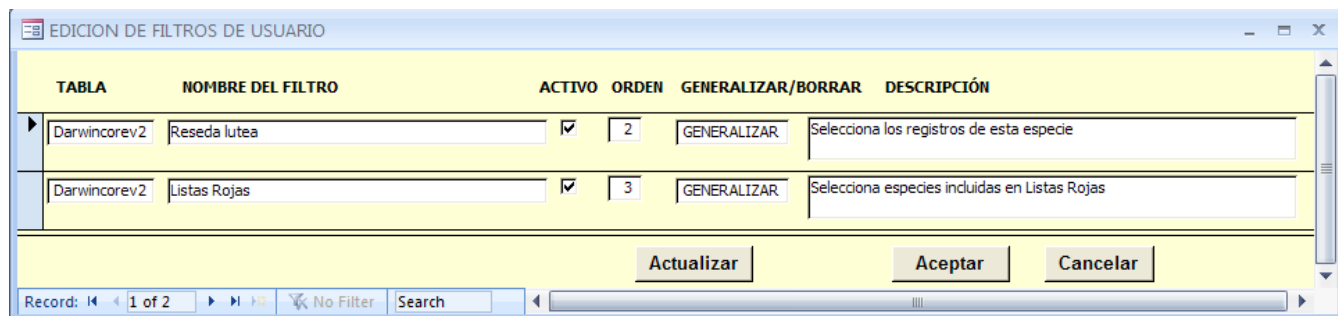


Figura 27.- Edición de filtros de usuario.

Para **aplicar alguno de los filtros** que hemos almacenado previamente, desde el formulario de selección de registros para generalizar pulsamos el botón **Seleccionar filtro**. Da paso a un formulario en el que aparecen todos los filtros creados por el usuario que permanecen como 'activados' en el formulario de gestión de filtros.

Basta con seleccionar el filtro deseado y pulsar **Seleccionar filtro**, para que las condiciones de selección se ejecuten.

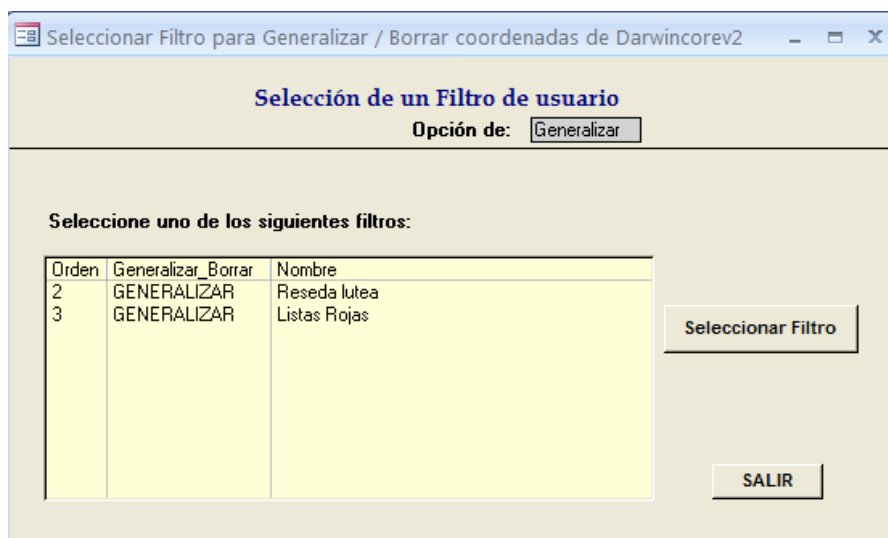


Figura 28.- Formulario de selección de filtros.

8. ELIMINAR REGISTROS DE DARWINCORE

Esta operación se realiza desde el botón **Borrar registros selectivamente de DwC vinculada** que abre un nuevo formulario.

Al igual que en el punto anterior de generalización o borrado de coordenadas, antes de realizar esta acción es conveniente realizar una copia de seguridad de la tabla **DARWINCORE**, para el caso de que los resultados no sean los deseados.

IMPORTANTE. NO debemos realizar **nunca** una copia de seguridad de una tabla **vinculada**, porque no es una verdadera copia, sino que se copia el acceso directo nada más y en realidad está apuntando a la misma tabla. Por ello, se debe hacer siempre la copia de seguridad de la tabla **DARWINCORE** en la base de datos en donde **reside realmente dicha tabla**.

Los registros que queremos eliminar se seleccionan eligiendo las condiciones de selección. Para limpiar las condiciones de la consulta que puedan haberse quedado almacenadas pulsamos el botón **Deseleccionarlos**.

Elegimos el campo por el que queremos filtrar los registros, el comparador y el valor que debe tener el campo elegido. Si queremos concatenar varias condiciones, entonces utilizaremos una de las opciones de campo UNION. En el ejemplo que se muestra a continuación, vamos a seleccionar los registros de la especie *Umbilicus rupestris* que queremos que sean eliminados de la tabla DARWINCORE. Una vez establecidas las condiciones de selección pulsamos el botón **Seleccionar registros**. Podemos ver la selección en forma de consulta, formulario o informe pulsando los botones **Ver consulta**, **Ver formulario** y **Ver informe**.

Figura 29.- Selección de condiciones de los registros que se van a excluir y borrar.

Si los registros seleccionados son los que deseamos eliminar, entonces pulsaremos el botón **Eliminar registros de Darwincore.**

Al igual que en el caso del formulario para difuminar o eliminar coordenadas, existe la posibilidad de almacenar y gestionar los filtros en este formulario. El funcionamiento es exactamente igual al explicado en el apartado [7. Difuminar o Eliminar de coordenadas de Darwincore.](#)

9. EXPORTACIÓN A DARWIN CORE ARCHIVE

Darwin Test incorpora una serie de procedimientos que permitirán exportar nuestra tabla DARWINCOREV2 o DARWINCORE14 al formato DARWIN CORE ARCHIVE (DwC-A). DwC-A es un nuevo formato estándar diseñado para simplificar y hacer más eficiente la publicación de los datos en la red del GBIF tanto a nivel de especies como de especímenes. Dicho archivo está compuesto por un archivo de texto *.txt* que contiene los datos, un archivo *eml.xml* que

contiene los metadatos o datos descriptivos de la colección, y un archivo *meta.xml*, que describe los campos con los que hemos trabajado en la exportación. Preferentemente estos ficheros se presentan comprimidos utilizando ZIP.



Figura 30.- Esquema del formato DwC Archive (http://www.gbif.org/orc/?doc_id=2816).

La generación del archivo DwC-A a partir de las tablas DARWINCOREV2 o DARWINCORE14 consta de dos pasos:

1. Cumplimentación del formulario con los metadatos de la colección al que se accede a través del botón **Metadatos de la colección**.
2. Exportación desde la opción **Exportar a Darwin Core Archive**.

9.1. METADATOS DE LA COLECCIÓN

Los metadatos son un elemento esencial de las colecciones que describen aspectos sobre los datos que contiene y la fuente que los proporciona. Desde el botón **Metadatos de la colección** se accede al formulario en el que deben cumplimentarse los metadatos para su futura exportación. Dicho formulario consta de 6 pestañas que recogen información básica sobre la colección, personal o institución proveedora, proyecto de investigación asociado, palabras claves y referencias relacionadas con la colección, así como otros metadatos adicionales. Los elementos de este formulario están definidos en el documento *GBIF Metadata Profile: Reference Guide* accesible desde el siguiente vínculo: http://www.gbif.org/orc/?doc_id=2820

IMPORTANTE. Todos los campos **en negrita** son obligatorios. Si dichos campos no se cumplimentan serán exportados con el valor N.A. (*Not Available*).

Metadatos básicos



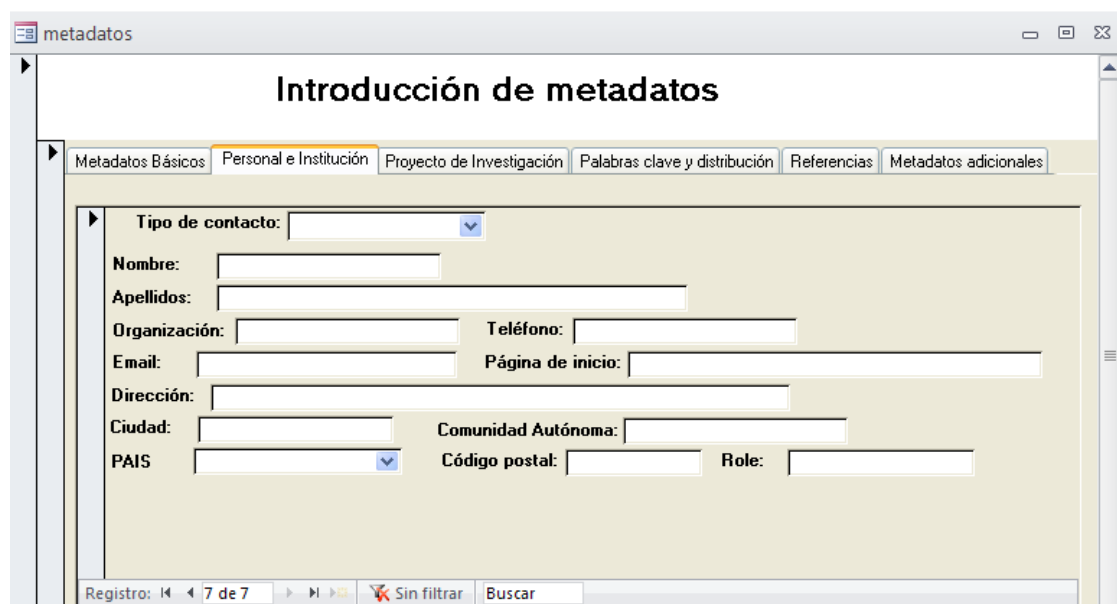
Figura 31.- Formulario de metadatos; metadatos básicos.

Esta pestaña recoge información básica sobre la colección repartida en los siguientes campos:

- **Identificador universal único:** código alfanumérico que identifica el documento EML. Puede ser obtenido desde siguiente generador <http://www.guidgenerator.com/>
- **Título:** título que describe la colección.
- **Fecha de publicación:** fecha de publicación de la colección en formato año-mes-día. Ej. 2010-09-20.
- Información adicional: información accesoria que el proveedor quiera incluir y que no esté recogida en ninguno de los elementos del formulario de metadatos.
- **Cómo citar:** cómo citar el recurso.
- Resumen: resumen de los datos que contiene la colección.
- **Idioma original:** idioma en el que están escritos los datos que contiene la colección, seleccionable desde un desplegable que recoge el estándar ISO 639-2.
- **URL:** URL del recurso si está disponible online.
- Licencia Creative Commons: licencia de los datos, a escoger de un desplegable entre las siguientes:
 - Public Domain
 - Creative Commons No copyright
 - Attribution

- Attribution non-commercial share-alike
 - Attribution share-alike
- Idioma Metadatos: idioma en el que están expresados los metadatos. La correspondiente casilla se rellena automáticamente tras cumplimentar los campos *Idioma* y *País*.

Personal e institución



The screenshot shows a web application window titled 'metadatos' with a sub-header 'Introducción de metadatos'. Below the header is a tabbed interface with the following tabs: 'Metadatos Básicos', 'Personal e Institución' (selected), 'Proyecto de Investigación', 'Palabras clave y distribución', 'Referencias', and 'Metadatos adicionales'. The 'Personal e Institución' tab contains a form with the following fields:

- Tipo de contacto: [dropdown menu]
- Nombre: [text input]
- Apellidos: [text input]
- Organización: [text input]
- Teléfono: [text input]
- Email: [text input]
- Página de inicio: [text input]
- Dirección: [text input]
- Ciudad: [text input]
- Comunidad Autónoma: [text input]
- PAIS: [dropdown menu]
- Código postal: [text input]
- Role: [text input]

At the bottom of the form, there is a navigation bar with the text 'Registro: 14 de 7', navigation arrows, a 'Sin filtrar' button, and a 'Buscar' button.

Figura 32.- Formulario de metadatos; personal e institución

Esta pestaña recoge información sobre la institución y el personal que proporciona la colección de datos. Para su correcta cumplimentación debe crearse un registro para cada tipo de contacto.

- **Tipo de contacto:**
- Creador del recurso: persona u organización responsable de la creación del juego de datos.
 - Proveedor de los metadatos: persona u organización responsable de la creación de los metadatos.
 - Contacto: institución o persona de contacto para cuestiones sobre el uso e interpretación del juego de datos.
 - Persona/institución asociada: otra persona u organización asociada con el recurso que interviene en la creación o mantenimiento del mismo.
 - Responsable del proyecto: persona u organización responsable del proyecto.

Para cada uno de estos tipos de contacto hay que rellenar los campos que siguen:

- **Nombre**
- **Apellidos**
- **Organización**
- **Teléfono**
- **Email**
- **Página de inicio**
- **Dirección**
- **Ciudad**
- **Comunidad autónoma**
- **País**
- **Código postal**
- **Role**

Proyecto de investigación

Este formulario recoge información sobre el proyecto de investigación en el que se incluye el juego de datos.

- **Título de la investigación:** título que describe el proyecto de investigación en el que se enmarca la colección de datos.
- **Descripción:** descripción del proyecto.

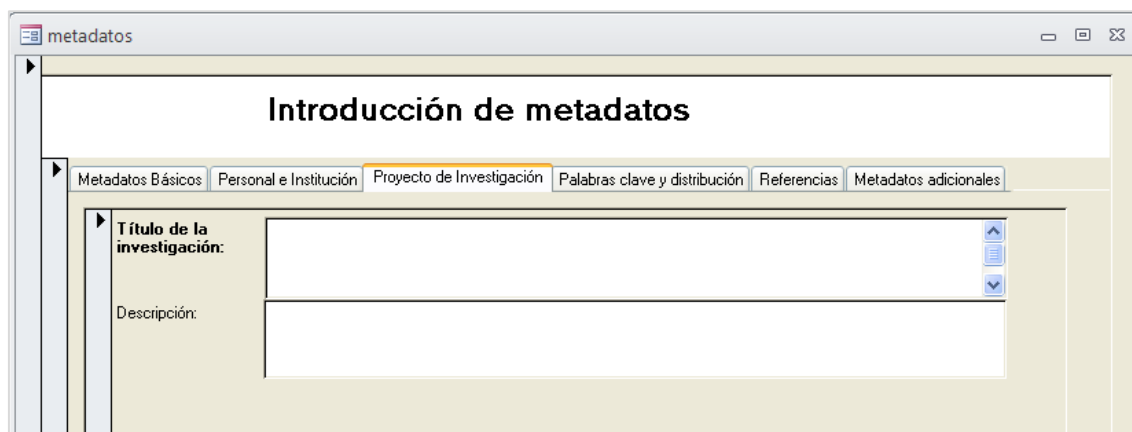


Figura 33.- Formulario de metadatos; proyecto de investigación.

Palabras clave y cobertura geográfica

Figura 34.- Formulario de metadatos; palabras clave y distribución.

Este formulario recoge información sobre la cobertura geográfica y palabras clave que identifican a los datos que contiene la colección.

- **Ámbito regional:** ámbito geográfico del proyecto.
- **Fecha de inicio:** fecha de inicio del proyecto con formato año-mes-día. Ej. 2010-09-20.
- **Fecha fin:** fecha de finalización del proyecto con formato año-mes-día. Ej. 2010-09-20.
- **Palabras clave**
 - **General**
 - Tesauros
 - **Taxonómicas**
- **Boundig Box:** se refiere a el área en forma de caja en la que se circunscriben geográficamente los datos de la colección (ver ejemplo de la figura 35 para Malta). Dicha caja está delimitada por dos valores de longitud y dos de latitud que se ingresan en los siguientes campos:
 - Boundig Box Oeste: valor de la longitud del margen izquierdo de la caja Ej. 14.180832
 - Boundig Box Este: valor de la longitud del margen derecho de la caja Ej. 14.57
 - Boundig Box Norte: valor de la latitud del margen superior de la caja Ej. 36.074997
 - Boundig Box Sur: valor de la latitud del margen inferior de la caja Ej. 35.799995

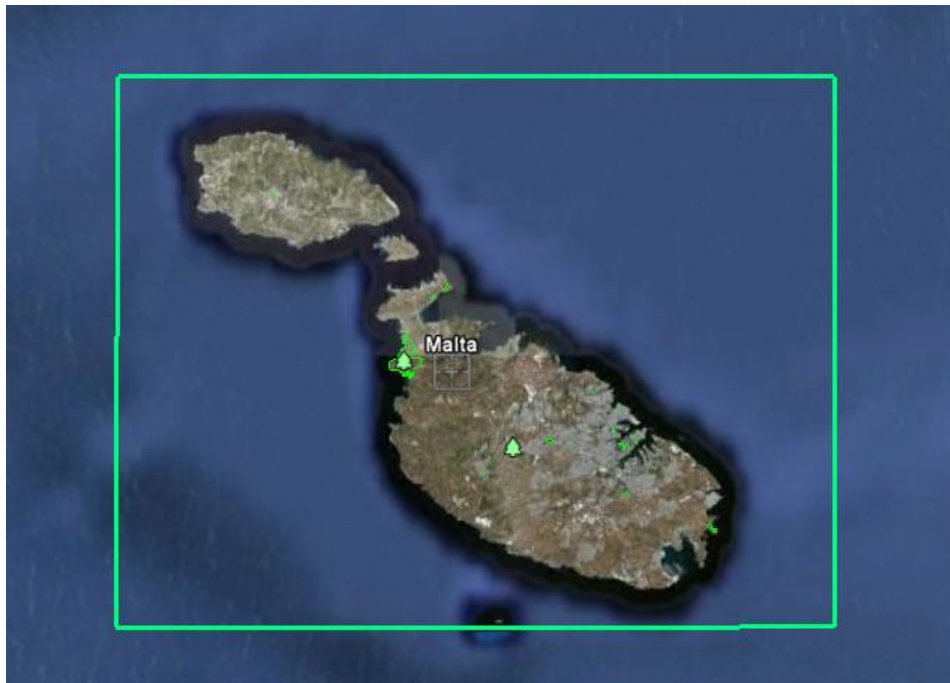


Figura 35.- Ejemplo de Boundig Box

Referencias

metadatos

Introducción de metadatos

Metadatos Básicos Personal e Institución Proyecto de Investigación Palabras clave y distribución **Referencias** Metadatos adicionales

Primera referencia:

Segunda referencia:

Tercera referencia:

Cuarta referencia:

Quinta referencia:

Figura 36.- Formulario de metadatos; referencias.

Esta pestaña recoge información sobre las referencias asociadas a la colección de datos.

Metadatos adicionales

The screenshot shows a web browser window titled 'metadatos'. The main heading is 'Introducción de metadatos'. Below the heading is a navigation bar with tabs: 'Metadatos Básicos', 'Personal e Institución', 'Proyecto de Investigación', 'Palabras clave y distribución', 'Referencias', and 'Metadatos adicionales' (which is highlighted). The form area contains the following fields:

- Fecha del sello: 03/08/2011 14:36:32
- Nivel Jerárquico: [input field]
- Tiempo de vida: [input field]
- url function: [input field]
- Collection Name: [input field]
- jgtiUnitType: [input field]
- Url del logo: [input field]
- Conservación del espécimen: [input field]
- Format Version: [input field]
- Parent collection: [input field]
- formation period: [input field]
- jgtiUnits: [input field]

Figura 37.- Formulario de metadatos; metadatos adicionales.

En este formulario se incluyen los siguientes campos:

- **Fecha de última modificación:** fecha en la que se han creado o modificado los metadatos adicionales.
- Nivel jerárquico: nivel al que se aplican los metadatos. *Dataset* es el valor por defecto.
- Tiempo geológico: para colecciones paleontológicas, el momento en el que el que vivió el organismo que se cita.
- URL: vínculo de interés adicional relacionado con la colección o institución que proporciona los datos.
- Nombre de la colección.
- Método de conservación: método de conservación del espécimen, a elegir entre un elemento de la lista desplegable con los valores disponibles en este vínculo: <http://rs.tdwg.org/ontology/voc/Collection#SpecimenPreservationMethodTypeTerm>
- jgtiUnit: número de unidades dentro de la colección. Es un campo de tipo numérico.
- jgtiUnitType: descripción general de la unidad de la colección (a lo que se pone número).
- Versión: versión del juego de datos (Ej. 1.0).
- Colección parental: colección general en la que se incluye la colección del juego de datos.
- Año de colección: texto que describe el periodo durante el cual tuvo lugar la campaña o campañas de colección comprendidas en el juego de datos.

9.2. CREACIÓN DEL ARCHIVO DARWIN CORE ARCHIVE

Desde el botón **Exportar a Darwin Core Archive** podemos obtener el archivo DwC-A en formato comprimido a partir de una tabla DARWINCORE. Antes de nada **es necesario haber vinculado la tabla DARWINCORE que nos interesa exportar.**

Al pulsar el botón la aplicación nos avisa del tiempo estimado que llevará el proceso de exportación que está directamente relacionado con el volumen de datos de la colección original. También nos recuerda que es imprescindible haber eliminado los caracteres ASCII anómalos previamente, ya que podrían dar problemas en la generación del archivo o en la indexación de los datos en el portal del GBIF, y nos da la posibilidad de eliminarlos en ese momento.

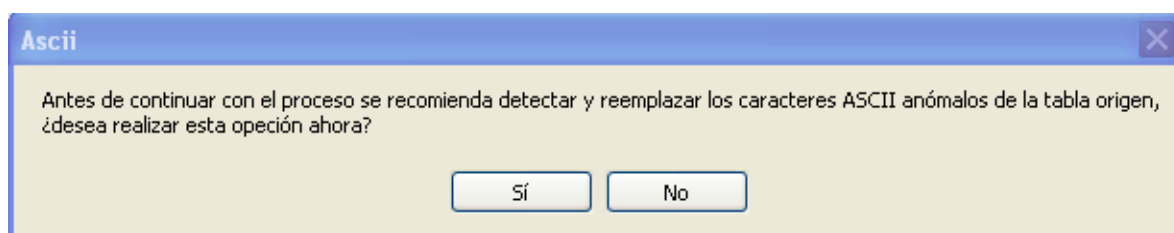


Figura 38.- Mensaje de aviso para la revisión de los caracteres ASCII.

Si pulsa **Sí** se pondrá en marcha el proceso que detecta y reemplaza los caracteres ASCII anómalos. Cuando realice este proceso se puede continuar con la exportación volviendo a pulsar el botón **Exportar a Darwin Core Archive.**

La exportación finaliza con un mensaje que nos informa de la ubicación del archivo generado. Por defecto, el fichero se almacena en C:/ con el nombre **DwcArchive.zip**

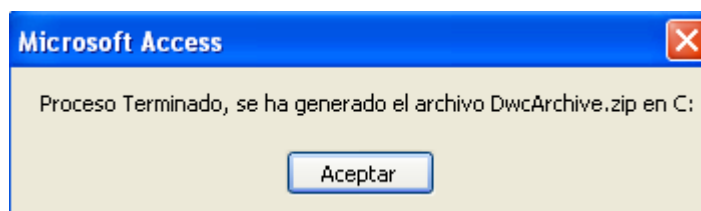


Figura 39.- Mensaje de finalización de la exportación

10. IMPORTAR ARCHIVO DARWIN CORE ARCHIVE

El botón **Importar archivo Darwin Core Archive** nos permite traer los datos en este formato a una tabla DARWINCOREV2 o DARWINCORE14.

Al pulsar este botón la aplicación nos recordará que debemos vincular una tabla Darwincore vacía (Darwincorev2 o Darwincore1.4, dependiendo del formato de origen) a la que se importarán los datos en formato Darwin Core Archive.

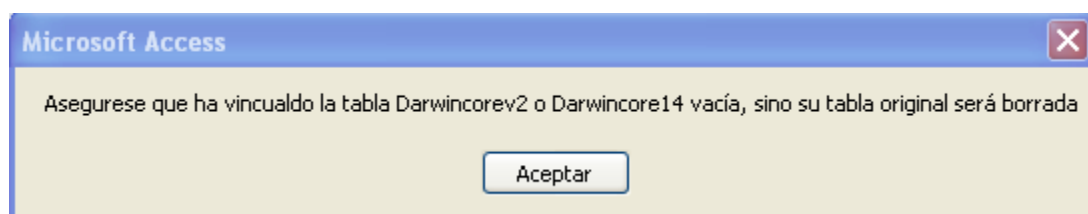


Figura 40.- Mensaje de aviso

A continuación la aplicación nos pedirá la ruta en la que se encuentra el archivo comprimido **DwCArchive.zip**. Una vez localizado utilizando el botón de navegación , hay que pulsar el botón **Generación del contenido de la tabla a partir del txt** para que finalice el proceso.

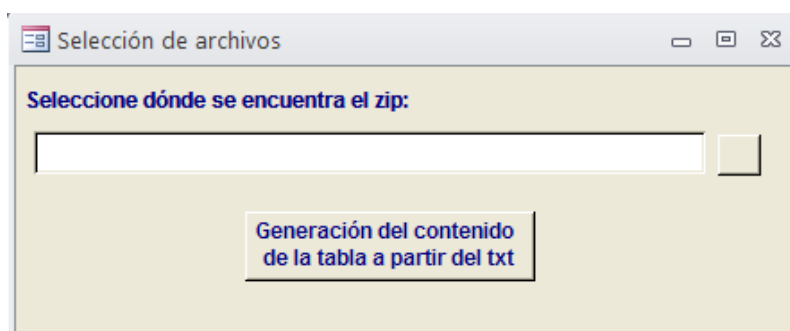


Figura 41.- Formulario para seleccionar la ruta con el archivo zip a importar.

Si los campos contenidos en el archivo Darwin Core Archive no correspondiesen con la tabla previamente vinculada (Darwincorev2 o Darwincore1.4) la aplicación mostrará el siguiente mensaje de aviso.



Figura 42.- Mensaje de aviso.

Una vez finalizado el proceso los datos se insertan en la tabla Darwin Core vinculada y aparece el siguiente mensaje de aviso.

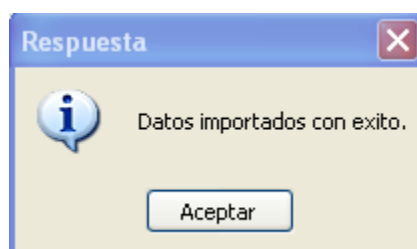


Figura 43.- Proceso de importación finalizado con éxito.

11. ÍNDICE DE CALIDAD APARENTE

Este índice se crea cada vez que validamos la tabla **DARWINCORE**. Nos indica la calidad **'Aparente'** de los registros de nuestra tabla y sirve para hacer un seguimiento de la mejora de la calidad en nuestros datos.

Conviene que cada tabla **DARWINCORE** tenga datos de una sola colección, ya que el Índice de Calidad Aparente se calcula por cada tabla completa de **DARWINCORE** y no por las posibles colecciones que ésta pudiera contener. En el caso de existir varias colecciones en la misma tabla **DARWINCORE** los datos del índice no serían válidos, ya que el índice se calcula respecto a la tabla completa y no a cada colección.

Los datos que se muestran en el Índice de Calidad Aparente son:

- **Total registros:** número total de registros que contiene la tabla **DARWINCORE** que hemos validado.
- **Total de registros con Especies/%:** número total de registros que vienen definidos al menos hasta nivel de especie. / El porcentaje que representan los registros con Especie respecto del número total de registros.
- **Total de registros con Países/%:** número total de registros que tienen definido el país. / El porcentaje que representan los registros con País respecto del número total de registros.

- **Total de registros Países OK/%:** número total de registros que tienen definido el país con código ISO. / El porcentaje que representan los registros con País OK respecto del número total de registros.
- **Total de registros con Coordenadas/%:** número total de registros que tienen definidas las coordenadas geográficas. / El porcentaje que representan los registros con Coordenadas geográficas respecto del número total de registros.
- **Total de registros con Coordenadas OK/%:** número total de registros que tienen definidas Coordenadas geográficas y éstas han superado la validación y están dentro del Boundingbox (rectángulo) del País y la Provincia. / El porcentaje que representan los registros con Coordenadas geográficas OK respecto del número total de registros.

Fecha validación	Colección	Total Registros	Total reg. con Especies %	Total reg. con Países %	Total reg. Países OK %	Total reg. con Coorden. %	Total reg. Coorden. OK %	Total reg. RadioPunto %
01/10/2009 10:53:25	Herbario_gbif	1,515	1,515 100	1,508 100	1,508 100	320 21	0 0	321 100
01/12/2009 14:54:11	Herbario_gbif	1,515	1,515 100	1,508 100	1,508 100	320 21	0 0	321 100
22/12/2009 10:30:12	Herbario_gbif	1,515	1,515 100	1,508 100	1,508 100	320 21	0 0	321 100
23/12/2009 12:27:59	Herbario_gbif	1,515	1,515 100	1,508 100	1,508 100	320 21	0 0	321 100
14/01/2010 9:27:15	Herbario_gbif	1,515	1,515 100	1,508 100	1,508 100	320 21	0 0	321 100
14/01/2010 11:23:08	Herbario_gbif	1,515	1,515 100	1,508 100	1,508 100	320 21	0 0	321 100
15/01/2010 8:26:22	Herbario_gbif	1,515	1,515 100	1,508 100	1,508 100	320 21	0 0	321 100

Tiene la opción de actualizar la información del Índice de Calidad Aparente en cada validación de Datos

Record: 1 of 11 No Filter Search

Figura 44.- Datos del índice de calidad aparente

Si quisiéramos borrar los datos del Índice de Calidad, pulsaremos la tecla **F11** y haremos click en **Tablas** de Access, seleccionaremos la tabla **ÍNDICE _ CALIDAD** y eliminaremos sus registros.

APÉNDICE I: TABLAS

Tablas de las bases de datos DARWIN_TEST. Algunas de las siguientes tablas aparecen duplicadas y una de las parejas tiene la terminación 14 (ejemplo, Filtro_difuminar / Filtro_difuminar14), lo cual se debe a que son tablas que tienen la misma función pero que se rellenan según si estamos utilizando la versión de Darwincorev2 o Darwincore14 de Darwincore.

TABLA	DESCRIPCIÓN
A_BASISOFRECORD	Todas las tablas que empiezan por 'A_' se generan de forma automática cada vez que se realiza la validación de datos. Si todos los datos están correctos, dichas tablas se borran automáticamente.
Babel	Tabla que almacena, en diferentes idiomas, los textos utilizados en los formularios de la aplicación.
CamposTablasRequeridos	Aquí se almacenan los campos que se van a modificar para las tablas que tengan problemas en el proceso de validación de datos.
CamposTablasValidados	Almacena los cambios efectuados por el usuario durante la validación.
CONGIF	Especifica el idioma actual en el que se encuentran los formularios.
CONTINENTES	Tabla de CONTINENTES y OCÉANOS estándar.
DARWINCOREV2/14	Tabla que debe contener los datos a validar.
Darwincorev2Validada/ Darwincorev14Validada	Tabla prácticamente gemela de DARWINCOREV pero que almacena los cambios realizados por el usuario durante la validación. Al contrario que la anterior no está vinculada y los cambios en ella no afectan a la tabla Darwincorev original. La diferencia de Darwincorev2Validada/ Darwincorev14Validada con Darwincorev es que esta primera tiene un campo de validación (sí/no) para cada campo, de tal forma que si el campo ha sido modificado su campo validado correspondiente aparecerá marcado.

Donde_config	<p>Tabla de gestión interna (no visible por el usuario).</p> <p>Guarda datos de la configuración de la aplicación: ruta o path donde se encuentra la base de datos con la tabla DARWINCOREV2 ruta o path donde se encuentra la base de datos con la tabla SCIENTIFIC_NAMES de Species2000 o la base externa que se desee. Valor de grado de generalización para las coordenadas de longitud y latitud de DARWINCOREV</p>
Filtro	<p>Tabla de gestión interna (no visible por el usuario).</p> <p>Guarda las condiciones para el filtro de los registros que se desean eliminar de DARWINCORE</p>
Filtro_difuminar	<p>Tabla de gestión interna.</p> <p>Guarda las condiciones para el filtro de los registros cuyas coordenadas se desea difuminar o borrar de DARWINCORE.</p>
Filtro_Eliminar	<p>Tabla de gestión interna.</p> <p>Guarda los registros seleccionados que van a eliminar de DARWINCORE</p>
ÍNDICE _ CALIDAD	<p>Contiene los datos de cada colección tras cada validación de datos, y se utiliza en el formulario de 'Índice de calidad aparente' para el cálculo de los porcentajes.</p>
ISO3166	<p>Contiene los códigos ISO de Países y los Boundingbox (rectángulos que encuadran) de cada país.</p>
ISO3166-2	<p>Contiene los códigos ISO de las provincias y los Boundingbox de las provincias españolas.</p>
REINOS	<p>Tabla que contiene los nombre de los Reinos estándar.</p>
Seleccion_borrar	<p>Tabla de gestión interna.</p> <p>Guarda los registros seleccionados que se desean eliminar de DARWINCORE</p>

Seleccion_difuminar	Tabla de gestión interna. Guarda los registros seleccionados cuyas coordenadas se desean borrar o difuminar en DARWINCORE
Sustituir_Ascii	Muestra el tipo de carácter ASCII extraño y el valor por el que será sustituido.
VALIDACIONES	Contiene todos los campos de DARWINCORE y los nombres de las consultas que validan cada campo. También contiene una casilla para saber si la validación está activada o no y la descripción de la validación.
Versión	Tabla de gestión interna. Guarda el número de versión de DARWIN_TEST.

Tabla 3.- Tablas de las bases de datos DARWIN_TEST.

APÉNDICE II: COMPACTAR Y REPARAR

Desde la versión 1.2 de DARWIN TEST se compacta y repara la aplicación cada vez que se cierra el programa. De todas formas, si necesitamos compactar y reparar la base de datos donde se encuentra NUESTRA TABLA DARWINCORE, será útil conocer cómo realizarlo.

Esta herramienta de MS Access es de gran utilidad en los siguientes casos:

- El tamaño de la base de datos es desproporcionado con la cantidad de datos que contiene y la velocidad de los procesos lenta.
- En caso de un corte de luz que pueda dañar la base de datos si ésta estaba siendo utilizada.

En ambos casos esta utilidad compacta y repara la base de datos, reorganizando la forma de almacenar la información, reduciendo el tamaño que ocupa y optimizando el acceso a los datos (mayor velocidad en los accesos a los datos).

Para compactar y reparar la base de dato, ésta se debe abrir de forma independiente y desde el menú de *MS-Access* pulsar las opciones siguientes según si utiliza Access 2003 o Access 2007:

ACCESS 2003

Herramientas/ Tools → Utilidades de Bases de datos/ DataBase Utilities → Compactar y reparar
Base de Datos/ Compact and Repair Database

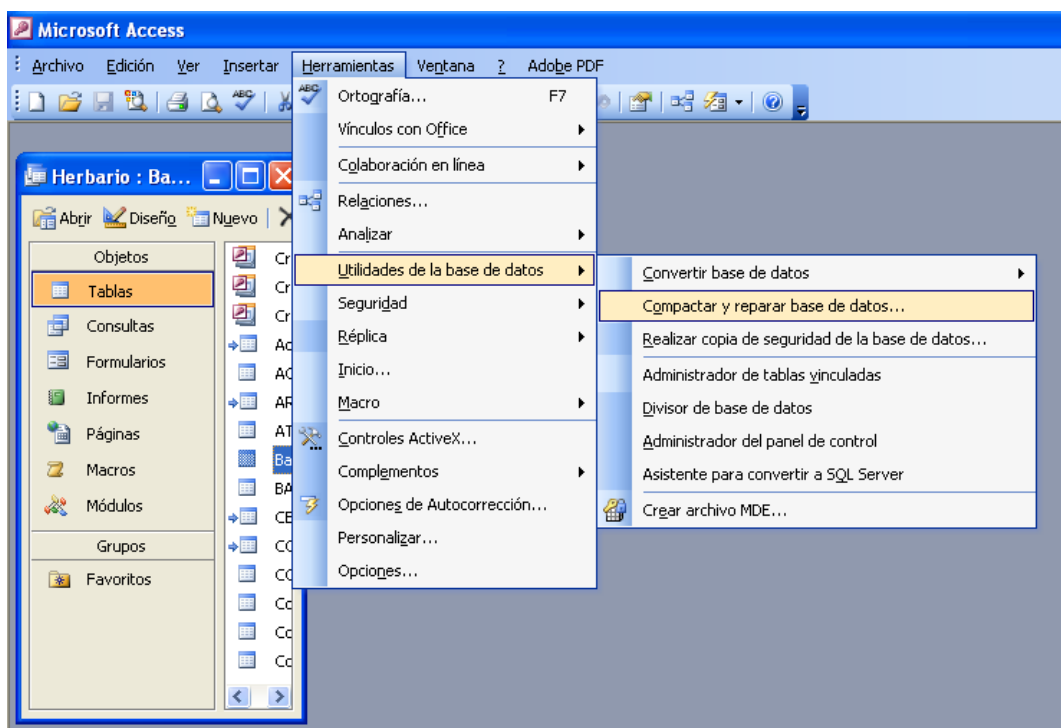


Figura 44.- Herramienta de *MS Access 2003* para compactar y reparar la base de datos

ACCESS 2007

Icono de Microsoft Office (esquina superior izquierda) → Gestión/ Manage → Compactar y
Reparar Base de datos/ Compact and Repair Database.

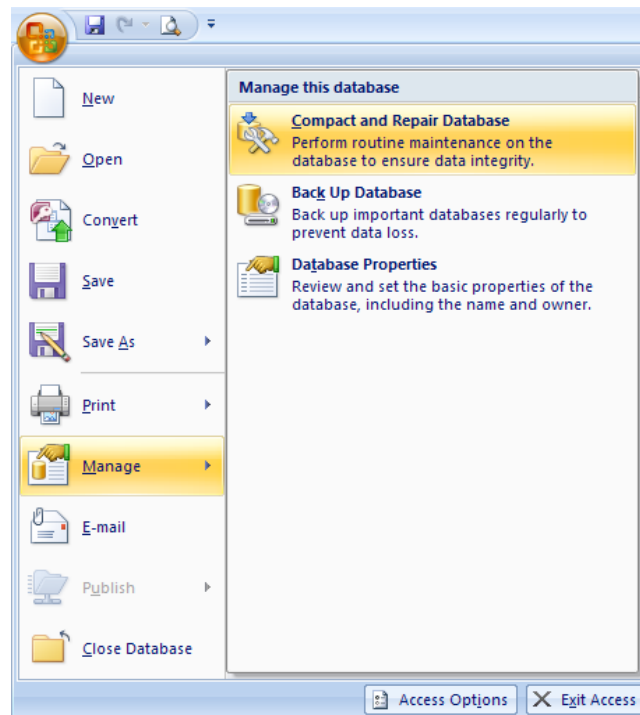


Figura 45.- Herramienta de *MS Access 2007* para compactar y reparar la base de datos.

ACCESS 2010

Archivo → Información → Compactar y reparar



Figura 46.- Herramienta de *MS Access 2010* para compactar y reparar la base de datos.

Darwin Test 3.1

Unidad de Coordinación GBIF.ES
2011