



# SiB

SISTEMA DE INFORMACIÓN  
SOBRE BIODIVERSIDAD  
DE COLOMBIA





# ESQUEMAS DE DIGITALIZACIÓN Y FLUJOS DE TRABAJO



# CONTENIDO

1. Introducción
2. Preguntas importantes
3. Grupos de tareas
4. Flujos de trabajo
5. Modelo de digitalización
6. Ayudas
7. Recomendaciones



# INTRODUCCIÓN

Lo más importante a la hora de iniciar un esfuerzo de digitalización de una colección biológica es definir un **MODELO DE DIGITALIZACIÓN** adecuado para la institución



# ALGUNAS PREGUNTAS IMPORTANTES

1. ¿Entiendo la importancia y el por qué de digitalizar una colección?
2. ¿Cuál es la dimensión de mis colecciones?
3. ¿Se digitalizará toda la colección? ¿Solo ejemplares tipo? ¿Representantes de cada grupo taxonómico?
3. ¿Datos e imágenes?
4. ¿Con qué recursos humanos y económicos cuento en este momento? ¿Tengo una proyección al futuro?
5. ¿Se puede hacer un estimado de cuánto puede constar la digitalización de la información asociada a un ejemplar? ¿\$0,5 USD 0 más de 10USD?



# GRUPOS DE TAREAS

ZooKeys 209: 19–45 (2012)  
doi: 10.3897/zookeys.209.3135  
www.zookeys.org

RESEARCH ARTICLE

A peer-reviewed open-access journal  
 ZooKeys  
Launched to accelerate biodiversity research

## Five task clusters that enable efficient and effective digitization of biological collections

Gil Nelson<sup>1</sup>, Deborah Paul<sup>1</sup>, Gregory Riccardi<sup>1</sup>, Austin R. Mast<sup>2</sup>

<sup>1</sup> *Institute for Digital Information, Florida State University, Tallahassee, FL 32306-2100, United States* <sup>2</sup> *Department of Biological Science, Florida State University, Tallahassee, FL 32306-4295, United States*

11 instituciones, 29 colecciones biológicas



# GRUPOS DE TAREAS ESPECÍFICAS

- Curado y pre-digitalización
- Captura de imágenes
- Procesamiento de imágenes
- Captura electrónica de los datos
- Georreferenciación



# CURADO Y PRE-DIGITALIZACIÓN

- Curado general de los ejemplares.
- Asignar un identificador único a cada unidad curatorial



# CAPTURA DE IMÁGENES

- Transporte de ejemplares desde la colección al lugar de fotografiado y viceversa
- Verificación de requerimientos técnicos de los equipos
- Fotografía de los ejemplares
- Paso de las imágenes a la estación de procesamiento



# PROCESAMIENTO DE IMÁGENES

- Nombramiento de las imágenes
- Control de calidad y retoma de imágenes
- Conversiones de formatos
- Recorte de imágenes
- Almacenamiento
- Copias de seguridad

# FOTOGRAFÍA DE TIPOS



Zootaxa 3357: 49–55 (2012)  
[www.mapress.com/zootaxa/](http://www.mapress.com/zootaxa/)

Copyright © 2012 · Magnolia Press

## Article

ISSN 1175-5326 (print edition)

**ZOOTAXA**

ISSN 1175-5334 (online edition)

**A new South American species of *Centris* (*Heterocentris*) Cockerell, 1899 with a key to the species with horn-like projections on the clypeus (Hymenoptera: Apidae: Centridini)**



CombineZP  
Software

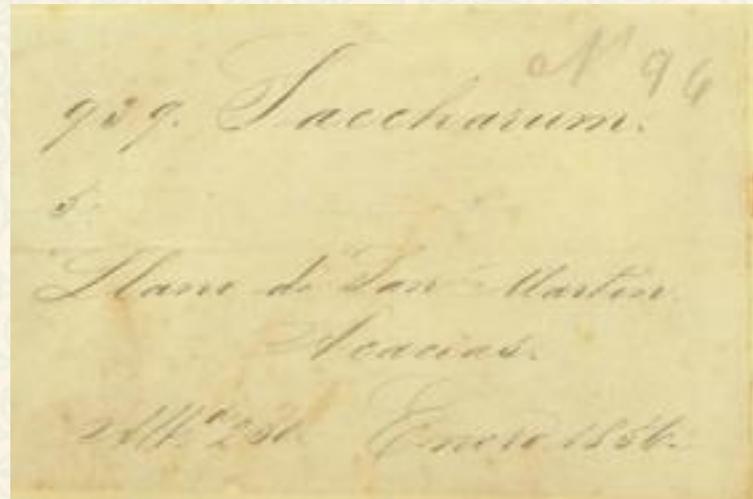


# CAPTURA ELECTRÓNICA DE LOS DATOS

- Captura convencional de datos con herramientas informáticas tradicionales
- *Optical Character Recognition (OCR)*
- Reconocimiento de voz

# CAPTURA ELECTRÓNICA DE LOS DATOS

- Optical Character Recognition (OCR)
- Tiene muchas ventajas comparativas con el método tradicional, en especial en ahorro de esfuerzo
- Es un poco complicado trabajar con textos antiguos escritos a mano.



# CAPTURA ELECTRÓNICA DE LOS DATOS

## Reconocimiento de voz

- Ya ha sido implementado en algunas colecciones de EEUU. Como parte de desarrollos privados.
- Existe una aplicación gratuita de Google Chrome para convertir voz a texto.

**Speech Recognizer**  
★★★★☆ (104) | [Productividad](#) | de [www.anilshanbhag.in](#) | 25.764 usuarios

**DESCRIPCIÓN GENERAL** | **DETALLES** | **OPINIONES** | **RELACIONADOS**

**Offline speech recognizer for chrome.**

This app uses the new built in speech feature in chrome 11+ to create a useful powerful tool to use voice to do the normal typing jobs.

The app is essentially an in the browser speech recognizer using which you can dictate sentences , edit the output and add them together to form paragraphs.

This app is best used along with Voice In chrome extension [ also developed by me ].  
<https://chrome.google.com/webstore/detail/gjkogbjkfcheifjonefnnehfggipnn>

Please Note :  
Speech Recognition is a statistical model . You cant expect a 100% accuracy rate . But not to worry recognition is rate is good enough to use :)

I'd appreciate if you'd rate it with ★★★★★ and tell your friends! :-)



# Speech Recognizer

col sucre coloso ser alto 29 0 1 10 col moreno y vargas



Spanish



col sucre coloso ser alto 29 0 1 10 col moreno y vargas

0).pdf - Adobe Reader

File Edit Documento Herramientas Ventana Ayuda

3 / 7 150% Buscar

pilosity. Primary pygidial plate with rounded apex, secondary plate with the distal margin truncated, as usual in the females of *Centris* (*Heterocentris*).

**Type material.** Holotype female with the following label: "COL. Sucre. Coloso. Ser. Alto 29-01-10 Col: Morelo & Vargas". The holotype is in good condition, but with the flagellum of the left antenna lacking. The type

# CAPTURA ELECTRÓNICA DE LOS DATOS

Métodos convencionales

<https://www.idigbio.org/content/biological-collections-databases>

Collections Management Databases

Product	Developer	Worksheet	Link to Public Comments (Forums)
BRAHMS V7 (Database/Portal)	Department of Plant Sciences Oxford UK	<a href="#">BRAHMS Worksheet</a>	<a href="#">BRAHMS Forum</a>
Electronic Museum (EMu) (Database)	KE Software Canada	<a href="#">EMu Worksheet</a>	<a href="#">EMu Forum</a>
Specify 6 (Database)	Biodiversity Research Center University of Kansas	<a href="#">Specify 6 Worksheet</a>	<a href="#">Specify 6 Forum</a>
Symbiota (Database/Portal)	Symbiota Software Project Global Institute of Sustainability (GIOS)	<a href="#">Symbiota Worksheet</a>	<a href="#">Symbiota Forum</a>

*Biodiversity Informatics*, 8, 2012, pp.1-11.

## A NEW ERA FOR SPECIMEN DATABASES AND BIODIVERSITY INFORMATION MANAGEMENT IN SOUTH AFRICA

WILLEM COETZER, OFER GON

*South African Institute for Aquatic Biodiversity, Somerset Street, Grahamstown, South Africa.*

*Email for correspondence: w.coetzer@saiab.ac.za.*

MICHELLE HAMER

*South African National Biodiversity Institute, Cussonia Ave, Pretoria, South Africa*

FATIMA PARKER-ALLIE

*South African National Biodiversity Institute, Rhodes Drive, Cape Town, South Africa*

*Abstract.* – We present observations and a commentary on the inherited legacy and current state of biodiversity information management in South African natural history museums, and make recommendations for the future. We emphasize the importance of using a recognized database application, and training and capacity development to improve the quality and integration of biodiversity information for research.

# CAPTURA ELECTRÓNICA DE LOS DATOS



*Ministerio de Ciencia, Tecnología  
e Innovación Productiva*



**SNDB**

Sistema Nacional de  
Datos Biológicos

"2010 - AÑO DEL BICENTENARIO DE LA REVOLUCIÓN DE MAYO"

CONICET



## Taller de evaluación de herramientas de captura y gestión de colecciones biológicas

Organizado por el "Proyecto de Fortalecimiento del Sistema de Digitalización e Integración de Registros Biológicos para la Conservación y Manejo de la Biodiversidad - Agencia de Cooperación Internacional del Japón JICA – CONICET", en conjunto con el SNDB

**INTA Castelar, 17 de agosto de 2010**

13 instituciones participaron en la evaluación



# CAPTURA ELECTRÓNICA DE LOS DATOS

## Conclusiones del taller

*“Herramientas recomendables. Sobre la base de la experiencia de los participantes se han detectado dos herramientas del ámbito internacional que surgen como buenos candidatos para implementarse en Argentina (**Zoorbar, Specify**). Ambas cuentan con versiones livianas que pueden utilizarse en colecciones pequeñas, y pueden migrarse a las aplicaciones principales, y exportan los datos a Darwin Core.”*



### HERBAR: Programa de gestión de colecciones botánicas



**HERBAR** es el nombre del **software para informatizar y gestionar colecciones botánicas** recomendado y apoyado por el nodo español de GBIF. Es un programa muy completo, con multitud de funciones y disponible en varios idiomas.

Es la aplicación estándar de la AHIM (Asociación de Herbarios Ibero-macaronésicos) y se utiliza de manera regular en decenas de instituciones españolas para la gestión de herbarios y bancos de germoplasma.

Para obtener más información sobre HERBAR, **visite su página web:**

<http://www.gbif.es/herbar/herbar.php>

### ZOORBAR: Programa de gestión de colecciones de historia natural



**ZOORBAR** es **software para informatizar y gestionar colecciones de historia natural** recomendado y apoyado por el nodo español de GBIF. Su flexible sistema de atributos permite su adaptación a las características particulares de cualquier colección biológica.

Esta aplicación está teniendo una interesante implantación en la comunidad zoológica española, y varias importantes instituciones ya gestionan sus colecciones de historia natural con ZOORBAR.

Para obtener más información sobre ZOORBAR, **visite su página web:**

<http://www.gbif.es/zoorbar/zoorbar.php>

### HERBAR/ZOORBAR LIGERO: Programa de gestión de colecciones botánicas y zoológicas



**Herbar-Zoorbar Ligero** es **una aplicación para informatizar y gestionar colecciones tanto botánicas como zoológicas**. Tiene como objetivo facilitar el intercambio de información entre los usuarios que informatizan el material biológico y la colección principal a través de tablas de Entrada rápida, sin necesidad de que los primeros se instalen la versión completa de Herbar 3.5 o Zoorbar.

Para obtener más información sobre HZL, **visite su página web:**

<http://www.gbif.es/hzl/hzl.php>



# GEORREFERENCIACIÓN

- Importante para la modelización datos
- Planeación de programas de muestreo y vacíos de información

The screenshot shows the GBIF website interface. At the top, there is a logo for GBIF (Global Biodiversity Information Facility) and the text "free and open access to biodiversity data". Below this, there are navigation buttons for "INFORMATICS", "PARTICIPATION", "GOVERNANCE", and "COMMUNICATIONS". A search bar is visible in the top right corner.

The main content area is titled "Online Resource Centre" and shows search results for the query "georeferencing". The search results are displayed in a list format, each with a thumbnail image, a title, a brief description, and a download count. The results are as follows:

Thumbnail	Title	Description	Downloads
	<b>Biogeomancer, Guide to Best Practices in Georeferencing</b>	A guide to the best practices for georeferencing biological species written by the BioGeomancer Consortium.	1763 downloads
	<b>MaNIS/HerpNet/ORNIS georeferencing guidelines</b>	Web document describing the georeferencing process in brief. Includes graphics, a glossary and a list of references. Last update from 2011.	35 downloads
	<b>Georeferencing for dummies</b>	A simple table that summarizes the MaNIS/HerpNet/ORNIS georeferencing guidelines. For a given locality type, it suggests what is the most common georeferencing procedure and how to determine the extent and error.	51 downloads
	<b>Presentation: Georeferencing concepts</b>	This presentation reviews the most important concepts around cartography, geographical information and the process of georeferencing.	48 downloads

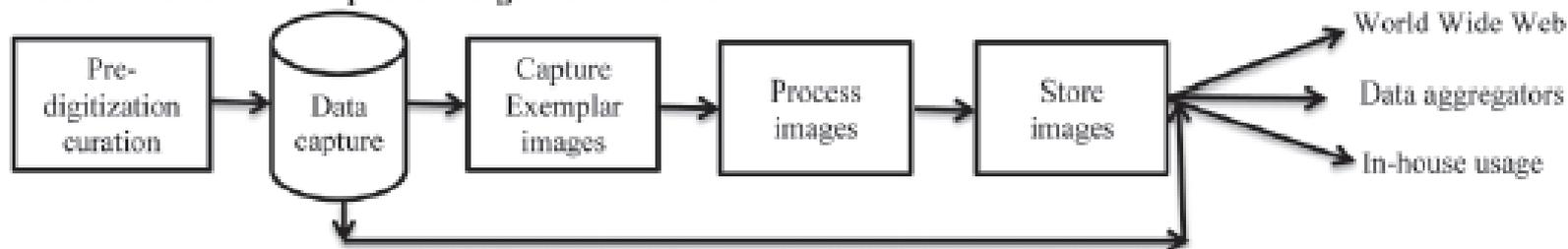


# GRUPOS DE TAREAS TRANSVERSALES

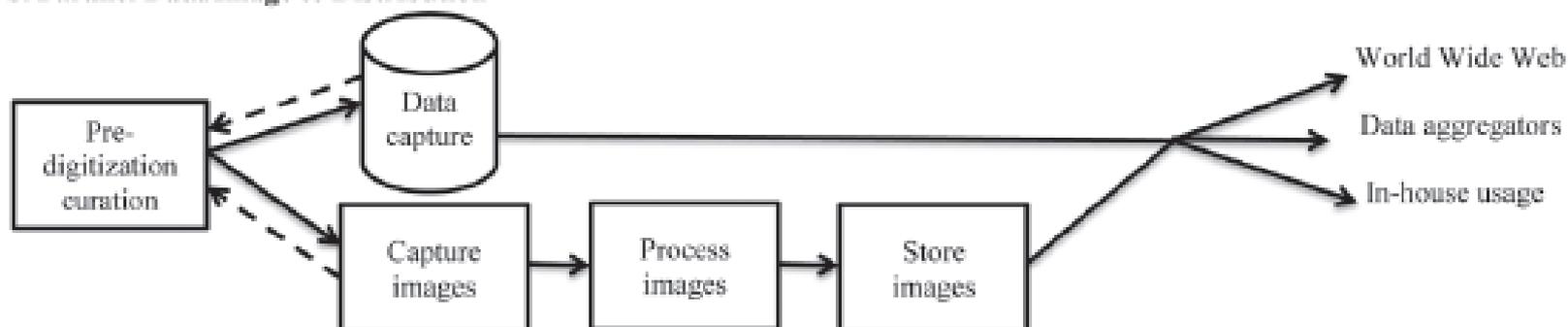
- Procesos de control de calidad
- Documentación de los procesos de digitalización
- Entrenamiento y capacitación

# ALGUNOS FLUJOS DE TRABAJO

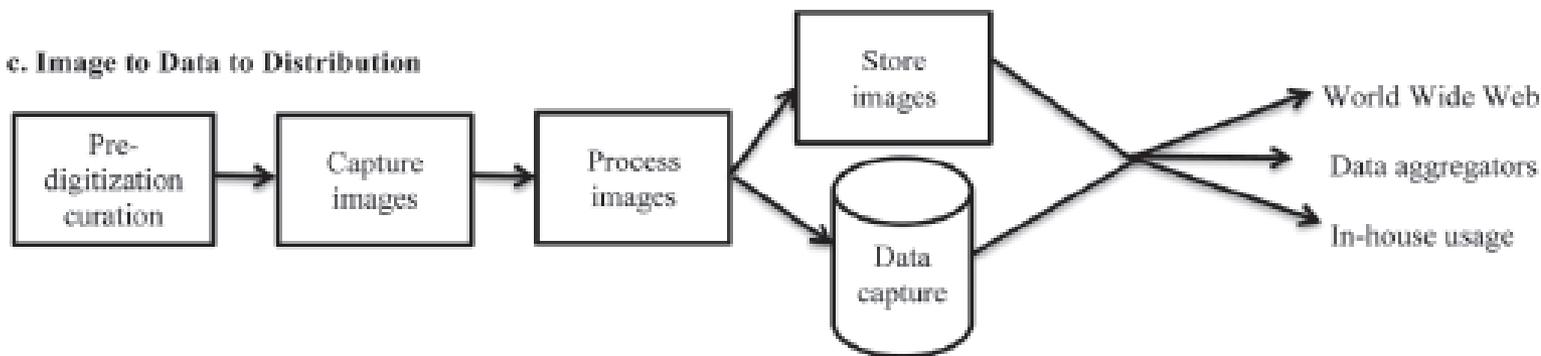
## a. Data to Occasional or Optional Image to Distribution



## b. Parallel Data/Image to Distribution



## c. Image to Data to Distribution





# MODELO DE DIGITALIZACIÓN

Una vez se tiene claridad en el por qué digitalizar, de los recursos con que se cuenta y de las tareas específicas a seguir se puede entrar a definir el modelo de digitalización más conveniente para la colección, esto debe involucrar aspectos importantes como:

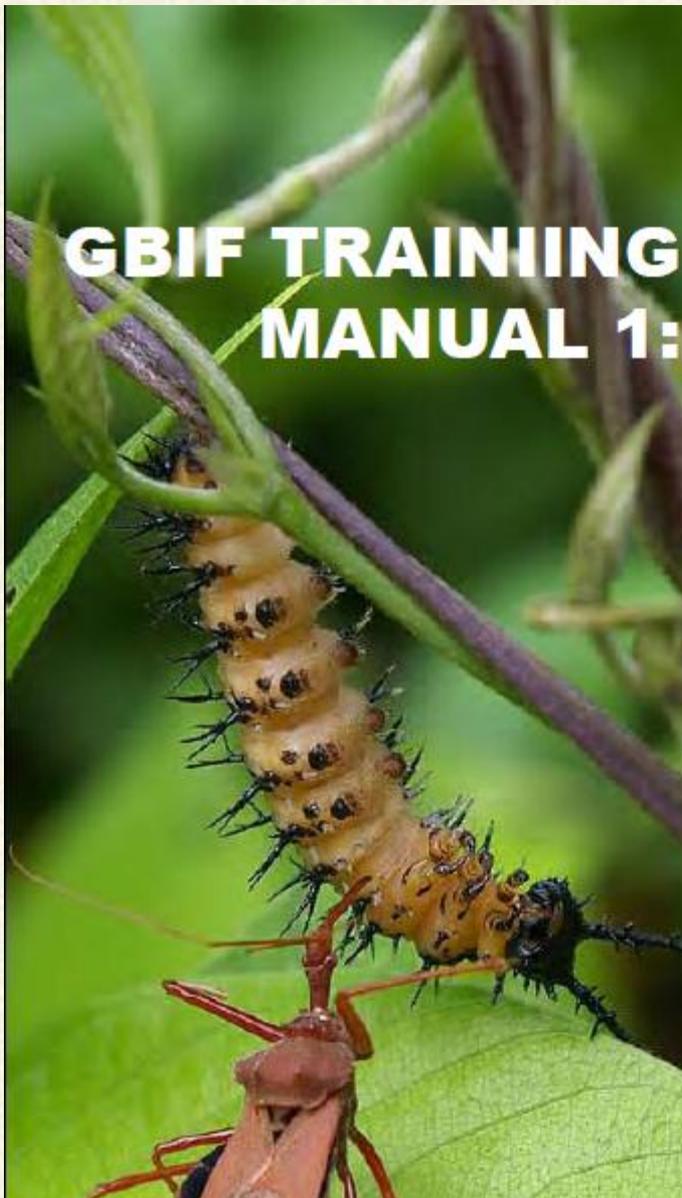
## 1. Definición de protocolos y manuales:

- Secuencia de pasos
- Captura de datos
- Toma y procesamiento de imágenes
- Almacenamiento de datos y media y
- Copias de seguridad
- Transporte de los ejemplares
- Control de calidad en cada proceso
- Actualización de identificaciones



# MODELO DE DIGITALIZACIÓN

2. Códigos de barras, QR e identificación de los ejemplares con información digitalizada dentro de la colección.
3. Definir sistemas de clasificación taxonómico y ordenamiento geográfico.
4. Definición del esquema de digitalización (centralizado o distribuido).
5. Definición del equipo humano y la arquitectura de trabajo que apoyará la digitalización de la colección.
6. Definir la herramienta informática a utilizar.
7. Definir el modelo de datos .



# **GBIF TRAINING MANUAL 1:**

## **DIGITISATION OF NATURAL HISTORY COLLECTIONS DATA**



# RECOMENDACIONES

1. Realizar un esfuerzo inicial importante en la planeación y definición del modelo de digitalización de la colección.
2. Realizar pruebas piloto antes de implementar de lleno el modelo.
3. Investigar a fondo las ventajas y desventajas de los software para la gestión de información de la colección.
4. Hacer revisiones periódicas del cumplimiento de protocolos de digitalización.
5. Realizar estimaciones del valor de un registro biológico en las condiciones de cada colección. Para apoyar la gestión asertiva de recursos.
6. Invertir tiempo y recursos periódicamente para la capacitación del personal involucrado en los procesos de digitalización de la colección.
7. Revisar con cierta regularidad los procesos y protocolos con el fin de sugerir mejoras que puedan ahorrar tiempo y dinero.



¡GRACIAS!  
OCTUBRE DE 2012  
TALLER DE COLECCIONES 3.0  
[sib@humboldt.org.co](mailto:sib@humboldt.org.co)