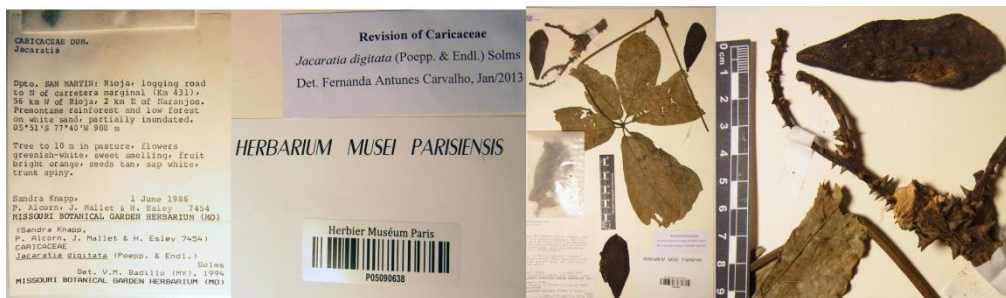


Using images to enter data in BRAHMS

Applies to v. 7.3 onwards

Fernanda Antunes Carvalho¹, Denis Filer² and Mike Hopkins³

1. Systematic Botany and Mycology, University of Munich (LMU), Munich, Germany
2. Department of Plant Sciences, University of Oxford, UK
3. INPA, Manaus, Brazil



Contents

INTRODUCTION	3
TAKING THE PICTURES.....	4
SPEED IMAGE AND DATA CAPTURE.....	4
FROM CAMERA TO WELL ORGANIZED FOLDERS.....	6
IMAGE ORIENTATION	7
A SUMMARY OF THE STEPS AHEAD.....	7
CREATE A NEW RDE FOR BOTANICAL RECORDS AND APPEND IMAGES	7
A NOTE ON MOVING IMAGE FOLDERS.....	8
TWO PHASES OF DATA ENTRY	8
ENTERING DATA TO RENAME IMAGES.....	8
ADDING A FIELD VIEW TO SPEED UP IMAGE RENAMING	10
THE POTENTIAL OF FOXPRO COMMANDS.....	11
ADDING SPECIES NAMES AT THIS STAGE	11
TYPING IN YOUR DATA	11
GENERATING NEW IMAGE FILE NAMES AND PHYSICALLY RE-NAMING IMAGE FILES	14
GATHERING IMAGES OF SAME-SPECIMEN RECORDS TO ONE RDE RECORD	16

Introduction

This document discusses how you can use ordinary images of specimens to speed up data entry. If your specimens are already digitized, you can adapt the methods here to link images to the existing data records.

We describe how you can quickly image herbarium specimens and their labels using a standard digital camera; register the images in a BRAHMS Rapid Data Entry (RDE) file for botanical records; enter sufficient data to allow image file renaming (vital for the future management of images); merge multiple images of the same specimen to one RDE data record; and enter the remaining data using an image of the label.

The advantages of imaging specimens extend to developing a database that combines data and images which in itself is useful for general database management, curation and, optionally, publishing online. Images bring a database to life! No special equipment is required except a digital camera.

This document assumes that you are already at least somewhat familiar with BRAHMS and some of the main methods and tools. If not, a useful way to start will be to explore the BRAHMS training guide available for download here:

<http://herbaria.plants.ox.ac.uk/bol/Content/Groups/brahms/Resources/BRAHMStraining2013.pdf>

The general BRAHMS training guide does refer to images and this tutorial is an extension to the section on Imaging specimens and their labels. It may also be useful to look at the section on FoxPro commands in the BRAHMS help file.

The method described here was used to build a database for the cyber-monograph of Caricaceae (<http://herbaria.plants.ox.ac.uk/bol/caricaceae>). It was initially tested by Mike Hopkins during the digitalization of the several herbaria in the Brazilian Amazon.

If, after reading this, you have any suggestions, problems, or general comments, please send these to antunesfc@gmail.com and/or to the BRAHMS project (advisory@brahmsonline.com).

Taking the pictures

The camera and basic setup

For taking images of herbarium specimens we suggest to use the Macro function (the camera used for the Caricaceae database was RICOH CX5) at resolutions of 7 megapixels. Do not forget a scale, so you can use the images later to measure morphological traits. If you are interested on morphometric analysis maybe it is interesting to use a tripod with a fixed distance to the specimen. However, this can also slow down the process, once you will need to adjust the zoom for taking pictures of labels, specimens and details. For the cyber-monograph of Caricaceae there was no tripod and the zoom and the basic camera set up were fixed.

Image resolution and purpose

If you want to image types (or other material) to provide something equivalent to what you can see using a binocular microscope on a physical specimen, you would need to image at a far higher resolution. However, for all curation purposes and many taxonomic ones too, smaller size images are perfect, far faster to take and more easily managed. If planning to image your entire herbarium, unlikely you will do all at high resolution.

If you are a taxonomist you may be interest on acquire a portable digital microscope. There are many models and brands with prices varying from 100 to 600 Euros. For the Caricaceae database I have a Dino-Lite AM-413ZT with the long-arm stand (MS 36B). The main advantages of a digital microscope is that it's light (c. 100g), easy to carry and allows we to take pictures easily of details such as flowers anther, trichomes, and others. Maybe it is more useful then images at a super high resolutions (which will need more external hard drives for storages and backups)

Speed image and data capture

One trained person can take *ca.* 1500 images per day – 750 specimens and another day to rename these images using the methods described here. Thus, a herbarium of 50,000 specimens could be completely imaged with data captured in *ca.* 6 months. If the data are already captured, the images can simply be linked to existing records.



The hand held approach offers max versatility and speed. The tripod offers more precision for further measurements of morphological characters using ImageJ, for example . Law of diminishing returns. A trade-off between time, cost and an end-product that is 'fit for purpose'.

Sequence of images on camera

Organizing your images is a process that must start when taking the photos

Does not matter if in the field or in the herbarium, the main information of a collection (collector number or label) should be your first photo.

In the field you need sequential photos from the same collection (collector number)

In the herbarium, you need sequential photos from the same specimen (accession number)

DSCN9672.JPG

DSCN9673.JPG

DSCN9674.JPG

DSCN9675.JPG

R0014681.JPG

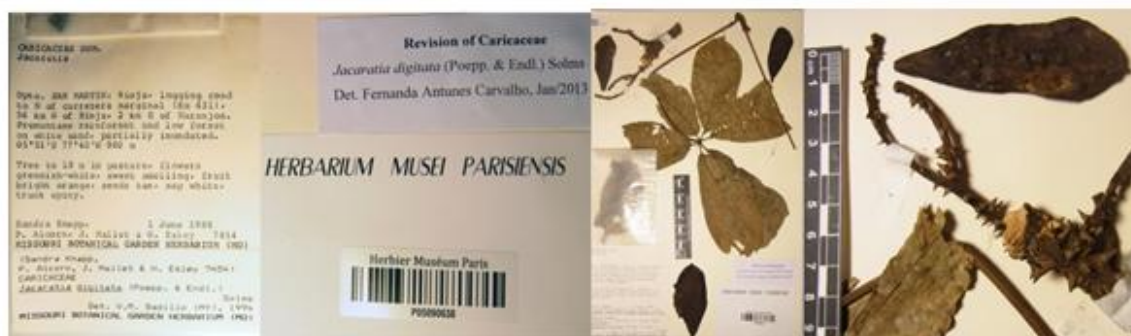
R0014682.JPG

R0014683.JPG

Important with managing the transfer of images from camera to disk to data file is image keeping track of the image sequence.

Images in the herbarium

For every herbarium sheet take 1st a picture of the label, this will normally include collector name, number, herbarium accession number, species name. Sometimes you may need a separate image for the barcode label. Then take photos of the specimen (complete herbarium sheet) and the details of your interest (flowers, fruits, venation, etc.).



DSCN00001.jpg

DSCN00002.jpg

DSCN00003.jpg

DSCN00004.jpg

A typical sequence of camera images from a single herbarium sheet

From camera to well organized folders

NB: Always best to make a backup of your original images.

Copy images from your camera to an organized folder structure. As with image file names themselves, any subfolders you create should not contain spaces or strange characters such as / , ; ! etc... Underscores ' _ ' are permitted. For the work on the cyber-monograph of Caricaceae, the lead author (FAC) is organizing her images in the following way:

- Herbaria
 - BM_2013feb
 - K_2013feb
 - MEXU_2012aug
- Expeditions
 - Venezuela_2012feb
 - Mexico_2012aug
 - Brazil_2012jan
- Living_plants
 - Jacaratia_spinosa_Janine_silva
 - Janilla_Jose_Lomeli
 - Vasconcellea_pubescens_Amanda_oliveira
 - Vasconcellea_pubescens_cultivated_Kew
 - Vasconcellea_microcarpa_cultivated_Munich
 - internet

Subfolder names already provide useful information about the images.

Image orientation

Images upside-down or otherwise badly rotated are quite annoying to view and do not look good if you plan to go online. Thus, before importing to BRAHMS, it's useful to have all images in the right direction.

To rotate many images at the same time in windows explorer select the images you need to rotate (ctrl + A to select all), right click on it and rotate all images according to your need (clockwise or counterclockwise)

A summary of the steps ahead

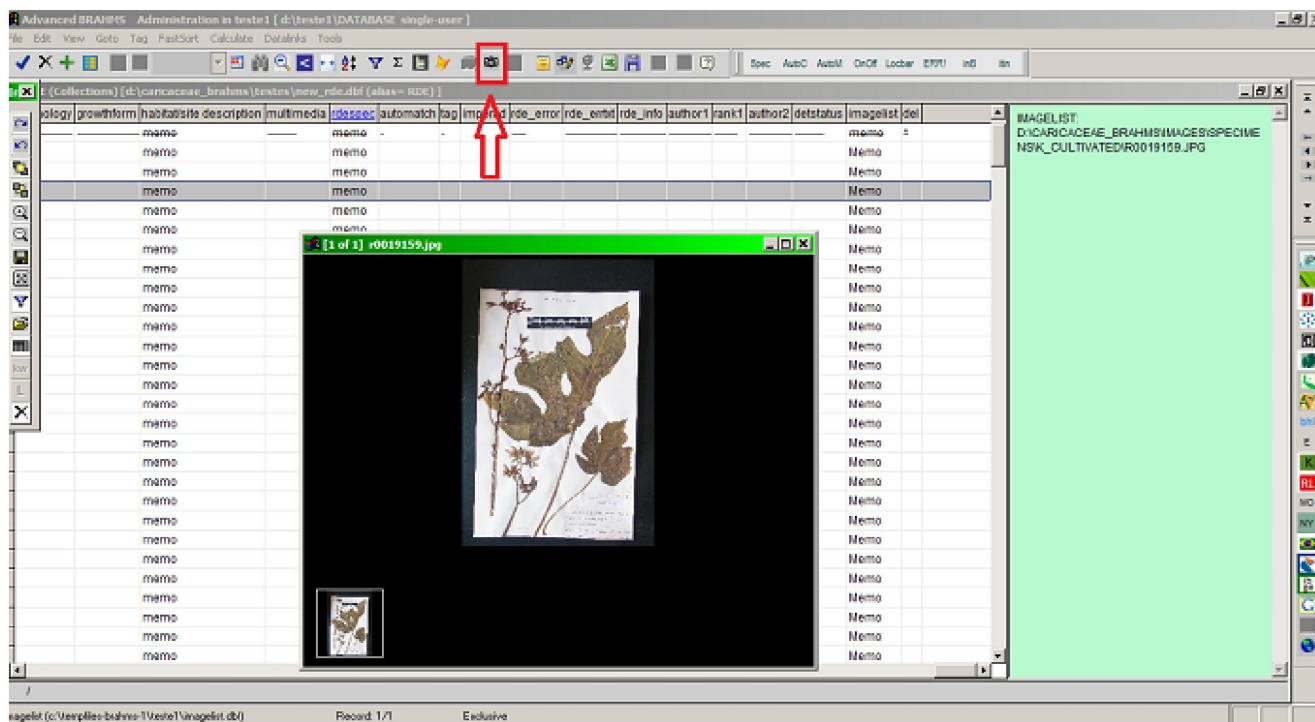
- Create a new RDE for botanical records and appending your images using the Tools option provided. This leads to one record per image in camera sequence.
- Typing into the RDE, from the images, the minimum data you need to rename your images files. The data you type in will be used to generate unique image file names – as you wish.
- Generate new image file names using the Tools options provided. This involves: a) generating a character string in IMAGETEMP based on the fields you choose; b) using IMAGETEMP to construct the full image path and file name based on the images in the IMAGELIST memo. These names are stored in the memo NEWNAME.
- Physically renaming your images on disk using the Tools option provided, based in the NEWNAMEs generated above. This process checks for potential errors such as duplicated file names.
- Merging records with multiple images per specimen. This leads to IMAGELIST memo having one to many entries.
- Using the images, typing in the remainder of the data.

Create a new RDE for Botanical Records and append images

- Select **BotanicalRecords > RDE file manager for Botanical Records > File > Create a new RDE** and follow the instructions. You may already have a project RDE template to copy but if not, you can use the BRAHMS system template.

Note that we don't use the option **Images > RDE file manager** and this may be a little confusing. That option is only for managing image files, not for entering specimen data.

- Open the new RDE file and then select **Tools > images in RDE > Append images to RDE files**. This adds one record per image, adding the image file name to the IMAGELIST memo field in the RDE file. Select any record and click on the camera icon  to see your image.



Note that when you add images like this, only the image path:file name is added to the file. Thus, the RDE file does not become large in size. You can see image file names by opening the IMAGELIST memo. For example: D:\CARICACEAE_BRAHMS\IMAGES\SPECIMENS\P\R0013408.JPG. As you will see later, you can store many images in the same memo field.

A note on moving image folders

If you subsequently decide to rename, move or otherwise re-organise your images folders, this is possible using the RDE option Tools > Images in RDE > Global editing of image drive:folder names to easily rename the drive:paths in your IMAGELIST memo.

If the images were already imported to database you can move images in the main Image library using tools< Re-locate tagged images to a new folder.

Two phases of data entry

- During the initial phase, you only need enter sufficient data to be able to correctly rename the images using the format you wish. And as discussed later, you will be able to merge records so that all images of the same specimen are linked to one record. In some cases, this may only be a barcode or accession number together with one further field that indicates what the image represents (specimen, label, etc.)
- During the second phase, you can enter the remaining data for the specimen - or as much of this as you wish. Examples will be date of collection, determination, location and perhaps some notes.

Entering data to rename images

Renaming images is vital – otherwise, images will always have the original camera file names – and these are not very helpful, especially as when renamed, you can use the image file name to auto-link images to data in BRAHMS.

For example, using the fields collector, number, dups, accession and/or barcode), you could easily rename images in one of the following ways, depending on your image naming policy where 188975 in these examples are an accession number:

DSCN00034.JPG → INPA_188975.JPG
DSCN00034.JPG → INPA_188975_Carvalho_2010_1_sd.JPG
DSCN00034.JPG → INPA-188975_label.JPG

Images can be renamed based on data you type into one or more fields. The idea is to type in the minimum data to allow a successful renaming of your images and thus merging of the records where there is more than one image per specimen (*i.e.* label(s) + specimen).

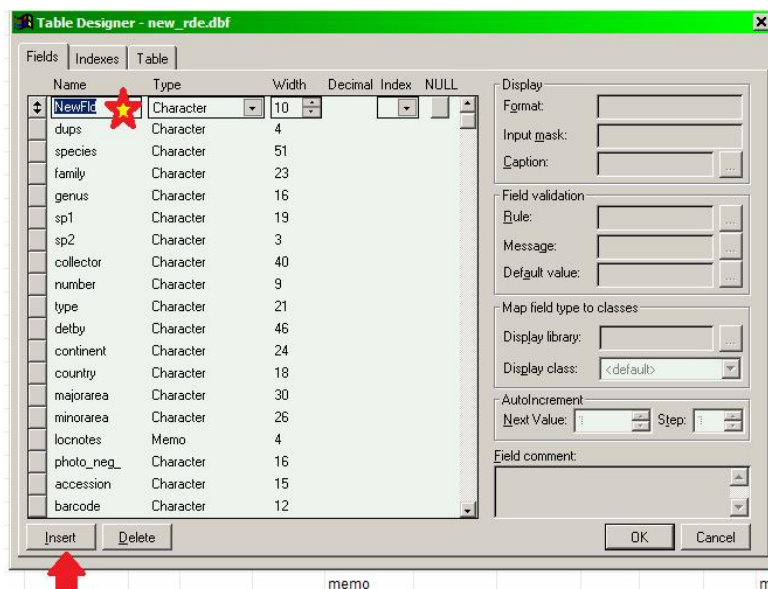
You can rename your images files in different ways and the naming system you use will depend on the available data and your preferences. Herbarium based projects tend to use barcodes or accession numbers for images names while research projects gathering data from multiple herbaria may take a different approach as you can see if the following example which includes the collector name and number as part of the image file name.

For the example that follows, first check you have the following fields in your RDE file – as these fields will be used to rename the images in this example.

DUPS
ACCESSION or BARCODE (depending on the availability)
COLLECTOR
NUMBER
PHOTO

If working in a herbarium digitization project, you could reduce this list to a minimum of BARCODE (or ACCESSION) and PHOTO as these could be sufficient to uniquely identify your specimen images.

The field PHOTO is used to differentiate images from the same collection/specimen. To create it go to **File > Modify file structure**. Click on insert and rename the new field to PHOTO.



Modifying the structure of an RDE file

You can create the field as character size 5 or larger if you plan to use longer photo suffixes. This field can be filled with a code you choose (e.g. L, L1 for labels, S, S1 for specimens or simply sequential numbers). I am using the following codes for my project:

- S = image of a specimen
- Sd = image of any detail of a specimen
- L = image of the collection label
- Lb = image of the barcode label
- Ld = image of the determinations labels

The PHOTO extension not only allows multiple images files with the same core name but it also facilitates searching and filtering later on. For example, you may want later to locate images of labels in a different folder and thus not Zoomify them for website use. Or you may want to use a default view in BRAHMS to show the label image by default when you open the image viewer.

INPA_12345_Antunes-Carvalho_FA_1234_l.jpg
 INPA_12345_Antunes-Carvalho_FA_1234_s.jpg
 INPA_12345_Antunes-Carvalho_FA_1234_1sd.jpg
 INPA_12345_Antunes-Carvalho_FA_1234_2sd.jpg

One could also think about more informative codes that may help to find an image according to the character. For example if you have of images of leaves coded as _leaf1, _leaf2, _leaf3 you could use a filter in the images main database file to see only leaf images of a species – although also see note on using image keywords below.

Adding a Field View to speed up image renaming

As the RDE file has many more fields than you need to see for this stage of data entry, it is useful to define a field view. At this point, I want to enter data in the following fields:

Dups, Barcode, Accession, Collector, Number, Photo

Optionally you may add to your Field View genus and sp1 as well (see below)

The potential of FoxPro commands

If all your records are from the same herbarium or collector you can use one or more Foxpro commands to fill those fields. For example:

Replace DUPS with 'M' all

Replace COLLECTOR with 'Carvalho, F.A.' all

Either now or after merging (see below), you may again use FoxPro commands to speed up general data entry, for example, if all from one country, Replace COUNTRY with 'Peru' all.

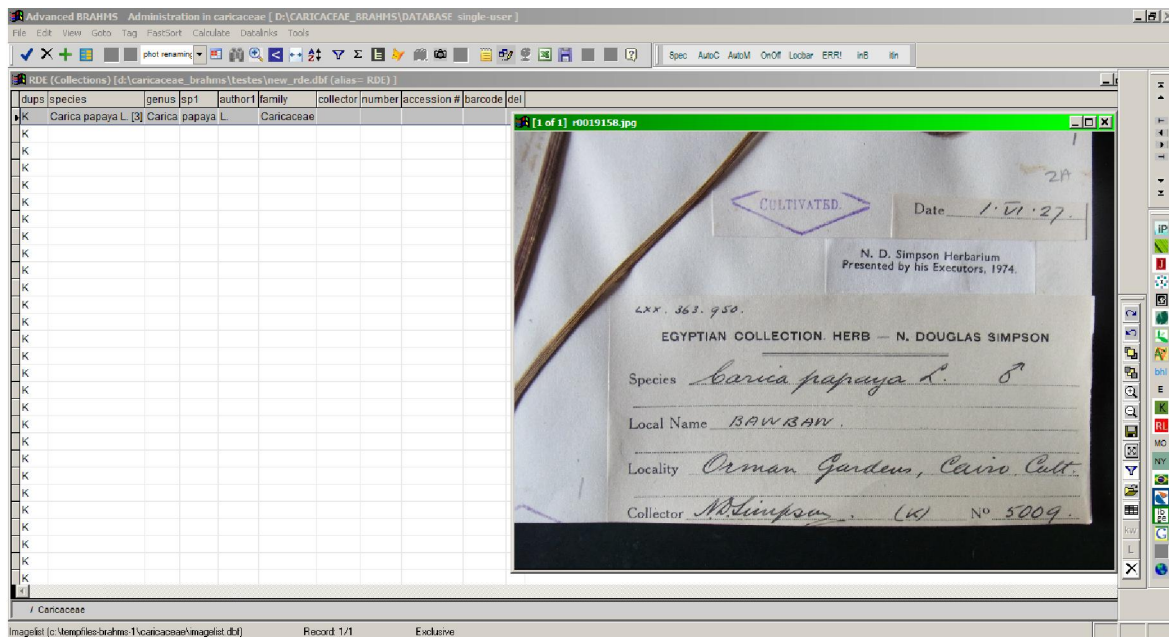
Adding species names at this stage

Although not essential at this stage, you may also want to include the species fields GENUS, SP1, AUTHOR1 in your field view. Some people prefer to enter the species name information at a later stage. But I (FAC) prefer to do it now because it increases the chance to detect identification/misspelling errors. Later on, you will open the images again and again to enter other kind of data, check descriptions, *etc.* and then you will be able to find misidentified specimens or errors you made while entering determination data. As soon as you have the determination fields filled, the higher the chance to find this kind of mistake.

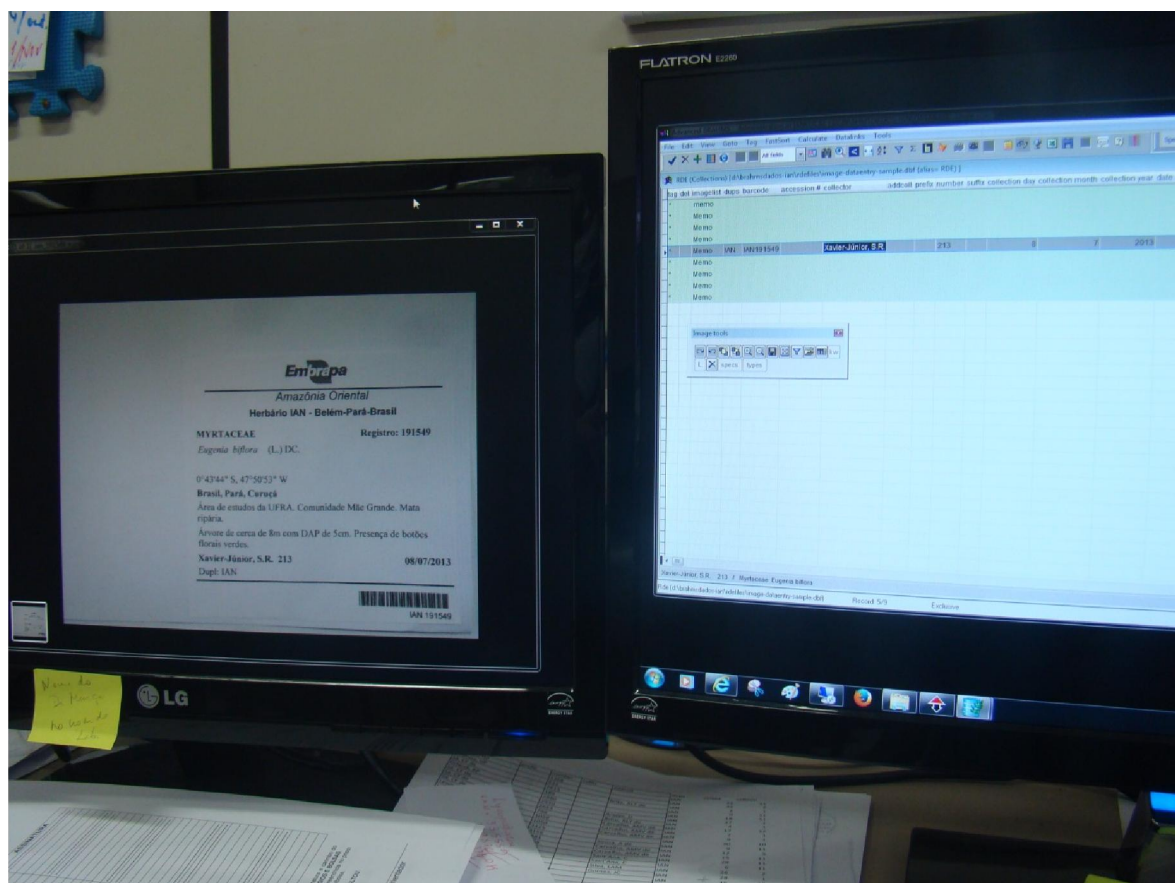
If you already have the species data entered in BRAHMS you can fill these three fields at once using F9 or a right-click in the SP1 field. It opens a window where you can search for a species name and when you select it the other fields are automatically filled. Then, you can copy down to all records for that specimen using F4 (see tips below)

Typing in your data

After configuring the field view, you can start to insert data in your database. Start from the 1st record that should be a label. With the image of the label open using image view, type the information in the visible fields.



Example of how the image of a label appears in a 13" screen with only some fields visible. You can locate the image view to the right and type the data in the columns to the left.



If you are fortunate to have a second monitor, drag the image viewer to the second screen. You can then enlarge the label image.

Some tips:

- Press F4 to copy information from a field in the record above.
- Press SHIFT + F4 to increment last field value by 1.
- If adding species names, right-click (or F9) on sp1 to search species name in the database.
- If adding collector names, fill collector in the BRAHMS format: Surname, initials as in Smith, A.B.
- Fill accession with 0 if there is neither accession number nor barcode. To differentiate different specimens from the same collection in the same herbarium use 0, 00, 000 in the accession number.

Advanced BRAHMS Administration in caricaceae [D:\CARICACEAE_BRAHMS\DATABASE single-user]										
File Edit View Goto Tag FastSort Calculate Datalinks Tools										
RDE (Collections) [d:\caricaceae_brahms\myrdefiles\images\k.dbf (alias= RDE)]										
dups	genus	sp1	author1	family	collector	number	accession #	barcode	del	
K	Vasconcellea	sphaerocarpa	(Garcia-Barr. & Hern.Cam.) V.M.Badillo	Caricaceae	Schunke, J.V.	5089				
K	Carica	papaya	L.	Caricaceae	Schäfer, P.A.	5589	0			
K	Carica	papaya	L.	Caricaceae	Schäfer, P.A.	5589	00			
K	Cylicomorpha	parviflora	Urb.	Caricaceae	Semsei, S.R.	3239				
K	Cylicomorpha	parviflora	Urb.	Caricaceae	Semsei, S.R.	4078				
K	Vasconcellea	quercifolia	A.St-Hil.	Caricaceae	Silva, F.C.	1836				
K	Carica	papaya	L.	Caricaceae	Sinclair	s.n.		000500512		
K	Carica	papaya	L.	Caricaceae	Sintenis, P.	312				
K	Jacaratia	dolichaula	(Donn.Sm.) Woodson	Caricaceae	Skutch, A.F.	2496				
K	Carica	papaya	L.	Caricaceae	Smith, H. H.	837				
K	Vasconcellea	glandulosa	A.DC.	Caricaceae	Smith, H.H.	838				
K	Cylicomorpha	parviflora	Urb.	Caricaceae	Smith, S.A.L.	288	000			
K	Cylicomorpha	parviflora	Urb.	Caricaceae	Smith, S.A.L.	288	00			
K	Cylicomorpha	parviflora	Urb.	Caricaceae	Smith, S.A.L.	288	0			
K	Jacaratia	spinosa	(Aubl.) A.DC.	Caricaceae	Solms-Laub.	2984	22626590			
K	Vasconcellea	cauliflora	A.DC.	Caricaceae	Soto, R.	4022				
K	Vasconcellea	glandulosa	A.DC.	Caricaceae	Spruce	s.n.				
K	Vasconcellea	parviflora	A.DC.	Caricaceae	Spruce	14924				
K	Vasconcellea	parviflora	A.DC.	Caricaceae	Spruce, R.	1865		000500527		
K	Vasconcellea	parviflora	A.DC.	Caricaceae	Spruce, R.	1865		000500528		
K	Vasconcellea	microcarpa	(Jacq.) A.DC.	Caricaceae	Spruce, R.	3859	0			
K	Vasconcellea	microcarpa	(Jacq.) A.DC.	Caricaceae	Spruce, R.	3859	00			
K	Vasconcellea	glandulosa	A.DC.	Caricaceae	Spruce, R.	4345				
K	Vasconcellea	glandulosa	A.DC.	Caricaceae	Spruce, R.	4345	20020039323			
K	Vasconcellea	glandulosa	A.DC.	Caricaceae	Steinbach, J.	7337				
K	Vasconcellea	quercifolia	A.St-Hil.	Caricaceae	Steinbach, J.	8637				
K	Vasconcellea	glandulosa	A.DC. ?	Caricaceae	Steyermark, J.	19237	00			
K	Vasconcellea	cauliflora	A.DC. ?	Caricaceae	Steyermark, J.	19237	0			

Example of a field view in a RDE of images after merging. Red boxes show specimens without accession/barcode number that were not gathered because of my (FAC) code to differentiate specimens from the same collection adding 0, 00, 000, etc. in the accession. Otherwise, all same-specimen images would have been merged correctly to a single specimen record. Note that the collection of Steyermark 19237, which has two specimens (herbarium sheets) in K, is identified with two different species names. This could be a mistake during data entering of sp1 or a real problem of two different species with the same collector number. In this case was a problem of data entry that I could detected when entering phenological data. The problem is even worst when the collector is unknown (quite common in K, for example)

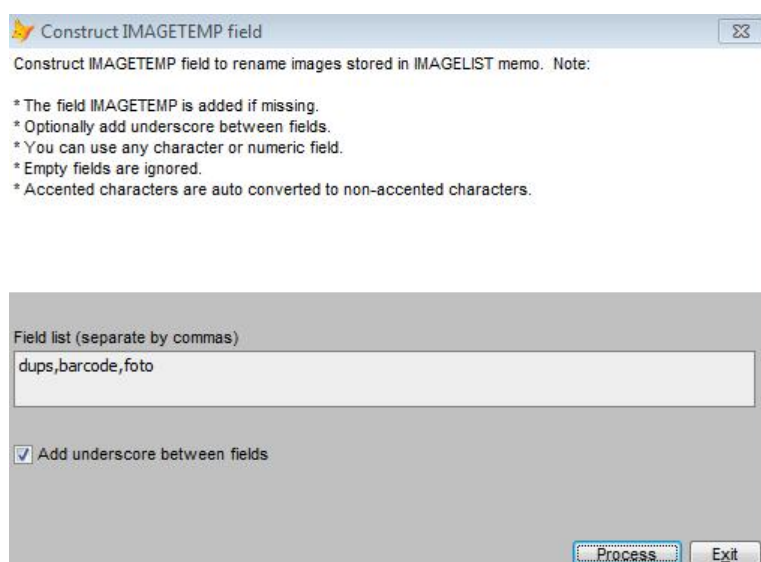
Generating new image file names and physically re-naming image files

After the fields above are filled for all records in your RDE you can generate the new file names and physically rename the images on your hard disk.

Step 1: Generate new file name strings

The new image files names are initially generated into the field IMAGETEMP. This name will not include the full path nor the file extension (e.g. .jpg), just a text string that will be used in a subsequent step (below) to create the new and completed image file names.

- Select **Tools> Images in RDE > Image file renaming and merging >Construct IMAGETEMP field for new image file names ...**



Enter a comma separated list of the fields from your RDE file you wish to use to build the new image file names and select the underscore option if you want that. This option creates the image file name text string based on these fields and adds this to IMAGETEMP. The field IMAGETEMP is added if missing from your RDE file. Certain accents such as é, ü, ô are auto-converted to non-accented format to avoid problems with illegal characters in image file names.

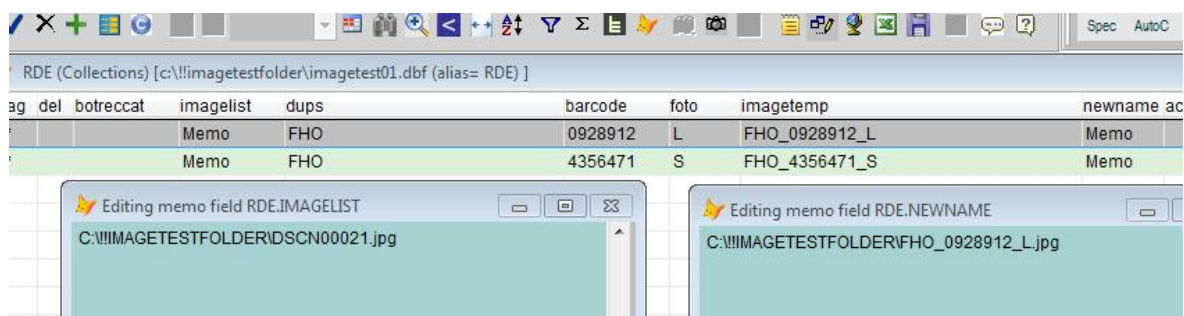


tag	del	botreccat	imagelist	dups	barcode	foto	imagetemp
*			Memo	FHO	0928912	L	FHO_0928912_L
*			Memo	FHO	4356471	S	FHO_4356471_S

The resulting string in IMAGETEMP. Note that, at this stage, the image path and image file extension (e.g. .jpg) are excluded.

Step 2: Generate NEWNAMEs

Once the strings are in IMAGETEMP, you can now choose Select **Tools> Images in RDE > Image file renaming and merging > IMAGETEMP -> NEWNAME for tagged**. This develops the full image name for each image in IMAGELIST using the text strings added above to IMAGETEMP.



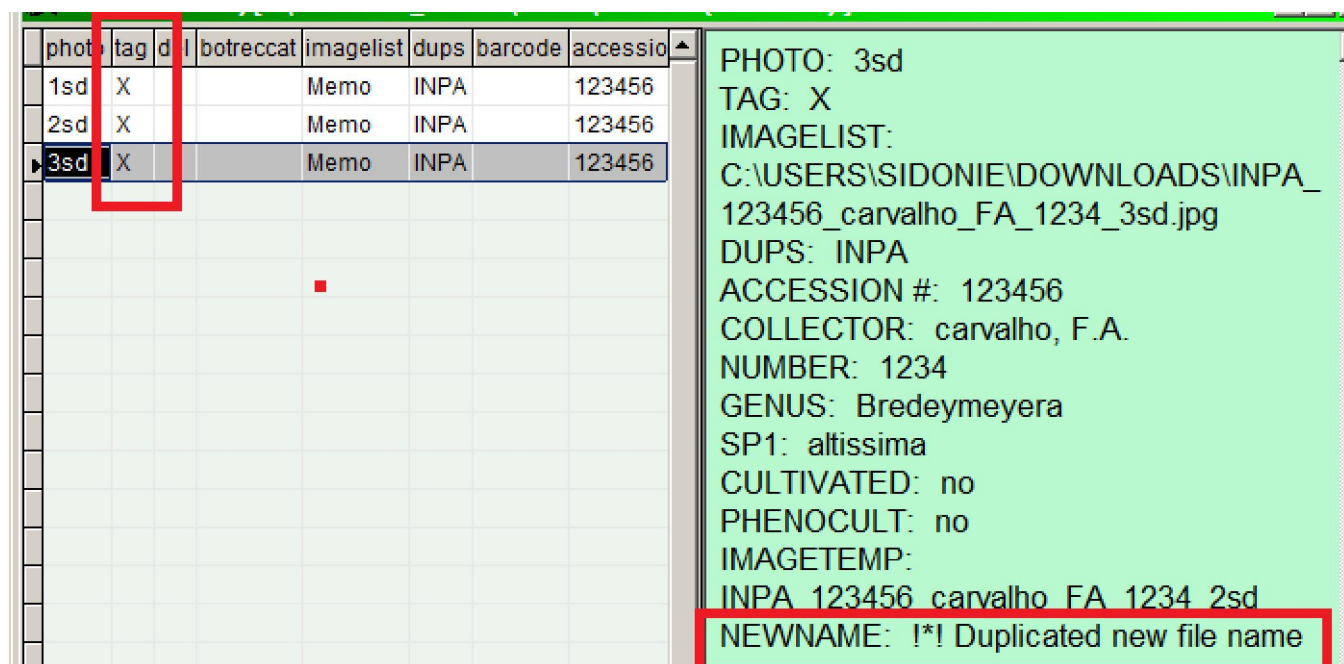
At this stage, the new names are ready to apply to the as yet un-edited names in IMAGELIST.

Step 3: Physically renaming images on disk

To finally update the image names in the IMAGELIST memo and physically rename the files on disk:

- Use **Tag > Tag all with *** to tag all records.
- Select **Tools > images in RDE > image file renaming > IMAGETEMP > - NEWNAME for tagged**

If there are records with identical file names, BRAHMS will tag these with X and add the message '!! Duplicated new file name' to the field NEWNAME.



You have to correct those names, usually by adjusting the field PHOTO. Repeat the process until you get rid of all tags= X. Then go to the next step.

- **Tools > images in RDE > image file renaming > physically rename tagged images files on disk to NEWNAMEs**

Now check your image folder. All files beautifully renamed!

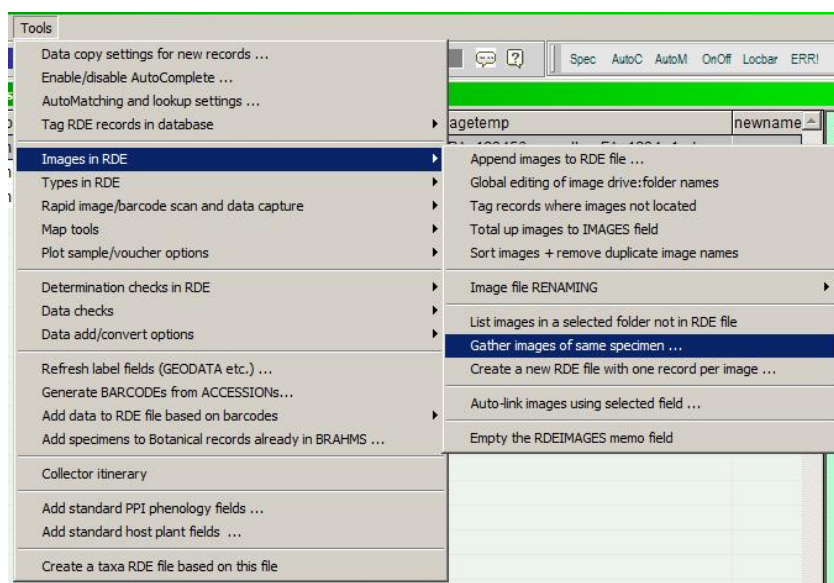
Gathering images of same-specimen records to one RDE record

Now you will join all images from the same specimens in one single record.

- For this, you will need a field that is the same for each specimen. If working in a herbarium digitization project you may use accession or barcode. However, when working in a taxonomic project (monographs, floras, etc.), with data from different herbaria this can be a problem because not all herbaria have accession or barcode numbers. A solution is to re-generate the IMAGETEMP field EXCLUDING the field PHOTO (which is used to differentiate each record/image). Select **Tools> Images in RDE > Image file renaming and merging >Construct IMAGETEMP field for new image file names ...**

Enter a comma separated list of the fields from your RDE file you used to build the image file names but this time, without the field PHOTO. This option creates an IMAGETEMP field that is the same for all records of the same specimens.

- In your RDE file, select **Tools > Images in RDE > Gather images of same specimen ...**



- Then select the fields you have to differentiate specimens (either accession or imagetemp as explained above)

Merge RDE images ✕

In those cases, where there are 2 or more images per specimen, these records can be merged to a single record, based on a selected RDE file field. Thus, if a specimen has 3 images in the RDE file (typically a specimen image and 2 separate label images), all 3 image file names will be added to one record - the other 2 records will be deleted.

Records are usually merged using accession number or barcode.

Use this option after renaming your camera image files.

Name field

imagetemp

▼

or accession, barcode

☒ Delete records that have been merged.

☒ Restrict to tagged

OK

Cancel