



Spreadsheets to expedite taxonomic publications by automatic generation of morphological descriptions and specimen lists

Visual guide, v1.5

by Ivan L.F. Magalhaes

División Aracnología, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" - CONICET .
Av. Ángel Gallardo 470, C1405DJR, Buenos Aires, Argentina. E-mail: magalhaes@macn.gov.ar

This manual explains the use of a group of three spreadsheets that are intended to help the writing of taxonomic works by automatically generating textual outputs for sections whose writing is generally tedious and time-consuming: the description of specimens, summaries of measurement variation and lists of examined material. Each of the spreadsheets is explained in detail in the following pages. Spreadsheets are protected to avoid that users accidentally modify formulas or other important content, but they can be unlocked and edited without the need for a password.

If using these spreadsheets, please cite: **Magalhaes ILF (2019) Spreadsheets to expedite taxonomic publications by automatic generation of morphological descriptions and specimen lists. Zootaxa 4624(1), 147–150. DOI: 10.11646/zootaxa.4624.1.12**

1. Descriptions

This spreadsheet outputs specimen descriptions. Each row corresponds to an individual description (e.g. the male of *Filistatinella hermosa*). Column A identifies the species. Columns C–F refer to sex/stage, type status, locality and collection number and all are printed in the beginning of the description. In the example below, the output from row 5 is “Male holotype from 32 miles E Laredo, Texas, USA (AMNH, IFM-1219).”

	A	B	C	D	E	F	G
1					Tags for description headers	<header>	Coloration
2					Tags for master characters	€€€	
3					Number of decimal digits for measurements:	2	
4	Species	Obs	Sex/stage	Type specimen	From	Voucher	
5	<i>Filistatinella crassipalpis</i>		Male	holotype	32 miles E Laredo, Texas, USA	AMNH, IFM-1219	
6	<i>Filistatinella crassipalpis</i>		Female		32 miles E Laredo, Texas, USA	AMNH, IFM-1220	as in male

Columns G onwards contain characters. The first row contains character headers, which can be used to separate your description into body regions; these headers are printed to the output. The second row contains *master characters*, which initiate new sentences in descriptions, and thus their name is also printed to the output. In the example below, the description of row 4 would read “Cephalothorax. Anterior margin of the carapace unmodified. Total length 1.66.” Please note that characters are only printed to the text of a particular species if the corresponding cell has data. In the example below, the description of row 4 would not mention its “Sternum”, since the corresponding cells lack data. **Important:** if there are multiple columns referring to the same header and master character, and at least one of them contains data, data should be inputted in the first column mentioning the header or master character, otherwise they will not be printed to the output. Also, columns referring to the same header or master character should be consecutive, otherwise they will be printed multiple times.

	A	U	V	W	X
1		Cephalothorax			
2		Anterior margin of the carapace	Sternum	Sternum	Total length
3					
4	Species		shape	sigillae	
5	<i>Filistatinella crassipalpis</i>	unmodified			1.66
6	<i>Filistatinella crassipalpis</i>	unmodified			1.6

	A	B	C	D	E	F	U	V	W	X	Y	Z
1					Tags for description headers	<header>	Cephalothorax					
2					Tags for master characters	€€€	Anterior margin of the carapace	Sternum	Sternum	Total length	Carapace	Carapace
3					Number of decimal digits for measurements:	2					length	width
4	Species	Obs	Sex/stage	Type specimen	From	Voucher		shape	sigillae			
5	Filistatinella crassipalpis		Male	holotype	32 miles E Laredo, Texas, USA	AMNH, IFM-12	unmodified	rounded	sigillae not visible	1.66	0.75	0.7
6	Filistatinella crassipalpis		Female		32 miles E Laredo, Texas, USA	AMNH, IFM-12	unmodified	rounded	sigillae not visible	1.6	0.7	0.62
7	Filistatinella pistrix		Male		San Bern. National Forest, Riverside, California, USA	AMNH, IFM-11	unmodified	rounded	sigillae not visible	2.33	0.99	0.93
8	Filistatinella pistrix		Female		San Bern. National Forest, Riverside, California, USA	AMNH, IFM-11	unmodified	rounded	sigillae not visible	2.60	1.04	0.90
9	Filistatinella hermosa		Male		Altadena, California, USA	MCZ 65568	unmodified	rounded	sigillae not visible	2.03	0.86	0.82
10	Filistatinella hermosa		Female		Santa Cruz Island, California, USA	AMNH, prep. IFM	unmodified	rounded	sigillae not visible	2.23	0.84	0.84

If several descriptors refer to the same master character, its name is printed only once at the beginning of the sentence.

Character headers (row 1) and master characters (row 2) can be tagged so they can be easily formatted latter. To tag them, introduce the corresponding tags into cells F1 and F2, respectively. Tagging is optional and these fields can be left blank. You can also use tags that describe format intuitively (e.g., <bold>, <ital>, <smallcaps>, etc).

If tagging is enabled, the output for row 5 in the example above would read “<header>Cephalothorax<header>. €€€Anterior margin of the carapace€€€ unmodified. €€€Sternum€€€ rounded, sigillae not visible. €€€Total length€€€ 1.66.”.

By applying the instructions explained below (see **Formatting the text using tags** below), one can use these tags to easily apply custom formatting to all descriptions in the manuscript. An example of formatted description would be applying bold + small caps to headers and italics to master characters:

CEPHALOTHORAX. *Anterior margin of the carapace* unmodified. *Sternum* rounded, sigillae not visible. *Total length* 1.66.

The third row contains *subordinate characters*. These are also printed to the text, but only continue sentences already started

	A	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX
1											
2		Leg I:	Leg I:	Leg I:	Leg I:	Leg I:	II:	II:	II:	II:	II:
3		femur (fe)	patella (pa)	tibia (ti)	metatarsus (mt)	tarsus (ta)	fe	pa	ti	mt	ta
4	Species										
5	Filistatinella crassipalpis	0.88	0.28	0.84	0.78	0.46	0.75	0.26	0.67	0.65	0.33
6	Filistatinella crassipalpis	0.74	0.25	0.65	0.55	0.41	0.6	0.22	0.5	0.47	0.33
7	Filistatinella pistris	1.14	0.32	1.03	1.03	0.58	1.01	0.33	0.93	0.98	0.46
8	Filistatinella pistris	0.97	0.32	0.88	0.82	0.52	0.85	0.33	0.74	0.75	0.43
9	Filistatinella hermosa	0.97	0.32	0.88	0.85	0.46	0.86	0.28	0.77	0.75	0.36
10	Filistatinella hermosa	0.86	0.32	0.77	0.71	0.46	0.72	0.26	0.61	0.6	0.38

The fourth row is not printed to the text; use it for comments, reminders, etc. The column headers are fully customizable, so change the character names to fit your purposes. All the characters are output in the order the columns are arranged. The current version of the spreadsheet supports 99 unique descriptors. If anyone should need to prepare descriptions with a larger number of characters, please let me know. **Important:** if you wish to add, remove, or rearrange the order of the characters, **do not** cut (ctrl+X) and paste data, and do not insert or delete columns. Use copy (ctrl+c) and paste to move data between columns G–DA and accommodate the characters. Cutting, inserting or deleting columns do not affect the formulas that concatenate the text, so if you make changes using these functions the reordering will not affect the output.

As you input the data, the descriptions are outputted in the *final description* sheet. Just copy the contents of the cells in column B (see below) and paste them into your favourite word processor. The formula trims and corrects the most common mistakes (double spaces, consecutive commas, etc).

[illegible]

2. Variation

This spreadsheet summarizes meristic counts and measurements. Each row should contain data for a single specimen. For the algorithm to work, the data should be sorted by species (column B) and then by sex/stage (column C). Row 3 contains character headers, which are included in the final output. The formula counts the number of measured individuals and fetches the minimum and maximum values of each character; outputting averages and standard deviations is optional (switch them on/off by typing “y” or “n” in the corresponding cells, A2 and C2). If a particular character is inputted in the table but you do not want it to be outputted to the final text (e.g. because you only needed it to calculate a ratio), you can prevent it from being outputted by typing “n” in the corresponding cell in row 2 (see example below: an “n” in J2 prevents “femur I length” from being outputted). Missing values are accepted, but the corresponding cell should have a textual string (such as a dash, -), otherwise it is interpreted as a zero.

	A	B	C	D	H	I	J	K	L
1	print mean values (Y/N)?	number of decimal values	print SD (Y/N)?	number of decimal values	Print?				
2	y	2	n	2	n				
3		species	sex/stage	voucher_identifier	total length	carapace length	femur I length	tibia I length	femur/carapace ratio
4		Filistatinella howdyall	Males		2.3	0.75	0.99	0.98	1.32
5		Filistatinella howdyall	Males		1.9	0.76	0.97	0.89	1.28
6		Filistatinella howdyall	Males		-	0.95	1.3	-	1.37
7		Filistatinella howdyall	Males		2.5	0.85	-	-	-
8		Filistatinella howdyall	Males		2.2	0.96	-	-	-
9		Filistatinella howdyall	Females		2.2	0.93	1.1	1	1.18
10		Filistatinella howdyall	Females		2.6	1	1.2	1.1	1.2
11		Filistatinella howdyall	Females		2.6	0.77	0.88	0.81	1.14
12		Filistatinella howdyall	Females		2.6	1	0.98	-	0.98
13		Filistatinella howdyall	Females		2.6	1.1	1.2	-	1.09
14		Filistatinella domestica	Males		2.4	1	1.3	1.3	1.3
15		Filistatinella domestica	Males		2.34	0.91	1.09	1.13	1.2
16		Filistatinella domestica	Males		2.77	1.15	1.54	1.5	1.34

The final text is outputted to the sheet *variation per species*; results are given separately by each sex/stage. Just copy the contents of the cells in column C (see below) and paste them into your favourite word processor.

	B	C
1	Species	Variation
2	Filistatinella howdyall	Males (n=5): total length 1.90–2.50 (2.23), carapace length 0.75–0.96 (0.85), tibia I length 0.89–0.98 (0.94), femur/carapace ratio 1.28–1.37 (1.32).
3	Filistatinella howdyall	Females (n=5): total length 2.20–2.60 (2.52), carapace length 0.77–1.10 (0.96), tibia I length 0.81–1.10 (0.97), femur/carapace ratio 0.98–1.20 (1.12).
4	Filistatinella domestica	Males (n=4): total length 2.34–2.77 (2.5), carapace length 0.91–1.15 (1.05), tibia I length 1.13–1.50 (1.32), femur/carapace ratio 1.20–1.34 (1.26).

3. Material examined

This spreadsheet generates lists of examined specimens sorted by locality. Data can be inputted in the following fields: **Collection number**, **Species**, **Type status**, **stage1**, **stage2**, **stage3**, **stage4**, **stage5**, **Country**, **Admin1**, **Admin2**, **Locality1**, **Locality2**, **Habitat**, **Collecting method**, **Altitude**, **Coordinates**, **Collector**, **Date**, **Obs**. Inputting data is only mandatory for the Species and Country fields (if there is no country data, I suggest inputting “No locality data” in this field); inputting data into Collection number and Admin1 fields is also strongly recommended. Other fields can be left blank and this will not affect the output. The headers of columns referring to stages (F–J) are printed to the text, so the user should rename them according to their needs (male, #f, ♀, immature, worker, etc.); these columns should contain the number of individuals.

	C	D	E	F	G	H	I	J	K	L	M	N
1	tag for country	tag for admin1	tag for admin2									
2	###	\$\$\$	%%%									
3	tag for type headers >>	&&&		Stage names								
4	Collection number	Species	Type status	#m	#ff	imm.	#x	#y	Country	Admin1	Admin2	Locality1
5	CAS 9057843	Filistatinella domestica		3	0	0			Mexico	Chiapas	Lagunas de Montebello	Posada de Bosque Bello
6	AMNH IFM-1209	Filistatinella domestica		0	1	0			Mexico	Estado de México	Almoloya de Juárez	Mina México
7	AMNH IFM-1208	Filistatinella domestica		1	0	0			Mexico	Estado de México	Almoloya de Juárez	Mina México
8	AMNH	Filistatinella domestica		0	2	0			Mexico	Hidalgo	2 miles NE Tizayuca	
9	AMNH	Filistatinella domestica		0	3	3			Mexico	Hidalgo	5 miles NW Actopan	
10	AMNH	Filistatinella domestica		0	2	2			Mexico	Hidalgo	5 miles S Zimapan	
11	AM KS.32582	Filistatinella domestica		1	1	0			Mexico	Hidalgo	bxmiquilpan	Río Tula
12	AMNH	Filistatinella domestica		0	2	0			Mexico	Hidalgo	bxmiquilpan	Río Tula
13	AMNH IFM-1211	Filistatinella domestica		11	10	0			Mexico	Hidalgo	bxmiquilpan	Río Tula
14	AMNH IFM-1158	Filistatinella domestica		1	1	0			Mexico	Hidalgo	Taxquillo (Tzindejeh)	
15	AMNH IFM-1163	Filistatinella domestica		0	1	0			Mexico	Oaxaca	9 miles SE Nochixtlan	
16	AMNH	Filistatinella domestica		0	1	0			Mexico	Puebla	Ozumbilla	
17	AMNH	Filistatinella domestica		0	1	0			Mexico	San Luis de Potosí	4 miles W San Luis de Potosí	

Before inputting the data, the *concat* sheet can be used to concatenate geographical coordinates and collecting dates into single columns. The data separator can be defined by the user. To concatenate the coordinates, the user should also include the brackets. While this adds an extra step, different records can have different brackets; I use this to differentiate between records whose original label included coordinates from those which I georeferenced myself. Including the altitude at this step is optional, as there is a separate column for it in the *data* sheet. **Important:** when copying data from the *concat* to the *data* sheets, remember to paste values (Ctrl + Alt + V, check the 'values' box, or right click > Paste special > values).

M	N	O	P	Q	R	S	T	U	V
Concatenate coordinates in the decimal format						Concatenate dates			
Altitude given in...: <input type="text" value="m"/>						Separator: <input type="text" value="/"/>			
bracket1	Lat-decimal	Long-decimal	Alt	bracket2	concat!	day	month	year	concat!
[22.319191	-102.167831]	[N22.31919°, W102.16783°]	30	VIII	1965	30/VIII/1965
[21.695222	-101.77337]	[N21.69522°, W101.77337°]	29	XII	1943	29/XII/1943
(19.49	-103.05)	(N19.49°, W103.05°)	10	V	1963	10/V/1963
(20.59	-104.02)	(N20.59°, W104.02°)	13	V	1963	13/V/1963
(19.57	-102.42)	(N19.57°, W102.42°)	11	V	1963	11/V/1963
(19.59	-102.41)	(N19.59°, W102.41°)	9	V	1963	9/V/1963
(19.70180556	-101.1995278)	(N19.70181°, W101.19953°)	20–25	VII	2014	20–25/VII/2014
(19.70180556	-101.1995278)	(N19.70181°, W101.19953°)	20–25	VII	2014	20–25/VII/2014
(19.70225	-101.1989444	1915)	(N19.70225°, W101.19894°, 1915m)	23	VII	2014	23/VII/2014
[19.70180556	-101.1995278]	[N19.70181°, W101.19953°]	20–25	VII	2014	20–25/VII/2014
(19.70180556	-101.1995278)	(N19.70181°, W101.19953°)	20–25	VII	2014	20–25/VII/2014
(19.70180556	-101.1995278)	(N19.70181°, W101.19953°)	20–25	VII	2014	20–25/VII/2014

The *convert* sheet can convert coordinates from degree-minutes-seconds to decimals and vice-versa. If you input degrees-minutes-seconds, please indicate the hemisphere using letters (N for north, S for south, E for east and W for west). When inputting coordinates as decimals, indicate the hemisphere by using positive (north, east) or negative (south, west) values. Failing to do so will result in the conversion returning an incorrect value.

Optionally, users can define the type status of the specimens. To do that, fill the following codes in the *Type status* (column E): ht = holotype, pt = paratype, lt = lectotype, nt = neotype, o = allotype, pl = paralectotype, st = syntype. Each category of type will have a separate list of examined material. Users can add new categories if they wish.

Other options for formatting the list are:

1) Locality names and type headers can be tagged using user-defined strings, such as \$\$\$, ###, <country>, <bold>, <i> etc. This allows the user to easily apply formatting to the list afterwards (see **Formatting the text using tags below**). Tagging is optional, so these fields can be left blank.

2) Users may define an abbreviation to be printed after collector's names (e.g. coll., leg.).

3) Users may define the punctuation mark after countries and first level administrative units using the "country separator" and "admin1 separator" cells

4) With the "verbose" option activated, the list will print "same collector", "same date" or "same collector and date" for some consecutive records with the same collecting data; this helps differentiating from instances where this information is lacking

d) The "pool" option will sum the number of conspecific individuals collected in the same locality and give them in a condensed format [e.g. 3#m (MACN-Ar 12768, UFMG 15968) instead of 2#m (MACN-Ar 12768), 1#m (UFMG 15968)]. When pooling with the "m" (minimal) option selected, museum numbers will be omitted [3#m (MACN-Ar, UFMG)]. For the minimal option to work, collection numbers must be separated from the museum acronym by a space.

	A	B
1	code	text
2	ht	Holotype:
3	lt	Lectotype:
4	nt	Neotype:
5	ot	Allotype:
6	pl	Paralectotype:
7	pt	Paratypes:
8	st	Syntypes:
9		
10		

tag for country	tag for admin1	tag for admin2	collector	verbose (y/n)	pool (n/y/m)				
<boldsmallicaps>	<bold>	<italics>	coll.	y	n			country separator	admin1 separator
tag for type headers >>	<boldsmallicaps>		Stage names					.	:

After all specimen and locality data has been inputted in the *data* sheet, it **must** be sorted in either of these two orders:

- 1) Species, Type status, Country, Admin1, Admin2, Locality1, Locality2, Habitat, Method, Coordinates, Collector, Date, Coll. Number
- 2) Species, Type status, Order-Country, Order-Admin1, Admin2, Locality1, Locality2, Habitat, Method, Coordinates, Collector, Date, Coll. Number

If you choose option 1, localities will be listed in alphabetical order. If you choose option 2, you can manually define the order in which countries and first-level administrative units are listed. To manually define this order, go to the *order-admin-units* sheet and enable manual ordering by entering “y” in cell I1. A list of your localities should appear in columns H and J (you might need to refresh formulas by hitting F9, and this might take a while). After they appear, define the order they should be listed in by inputting numbers in the columns I and K and sort your data using option 2 (see above).

	H	I	J	K
1	Order manually? (y/n)	y		
2	Country	User defined order	Admin1	User defined order
3	Mexico	2	Mexico@Aguas Calientes	3
4	United States of America	1	Mexico@Chiapas	12
5			Mexico@Estado de México	8
6			Mexico@Hidalgo	7
7			Mexico@Jalisco	4
8			Mexico@Michoacán	5
9			Mexico@Nayarit	1
10			Mexico@Oaxaca	11
11			Mexico@Puebla	10
12			Mexico@San Luis de Potosí	6
13			Mexico@Tamaulipas	9
14			Mexico@Zacatecas	2
15			United States of America@Arizona	2
16			United States of America@California	1
17			United States of America@New Mexico	3
18			United States of America@Texas	4

Important: remember to sort your data, otherwise the lists will be printed incorrectly! The lists are outputted to the Ready! sheet and given by species. Just copy the contents of the cells in column C (see below) and paste them into your favourite word processor.

	B	C
1	Species	Material examined (copy these cells and paste into text)
2	<i>Filistatinella domestica</i>	&&&Non-type material: &&&###Mexico.### \$\$\$Chiapas:\$\$\$ %%%Lagunas de Montebello,%% Posada de Bosque Bello, [E22.31919°, S102.16783°], P.R. Cra
3	<i>Filistatinella hermosa</i>	&&&Holotype: &&&###United States of America.### \$\$\$California:\$\$\$ %%%Los Angeles,%% Altadena, (E19.40954°, S99.72658°, 2618m), L. Pinter, 22/II/20
4	<i>Filistatinella kahloae</i>	&&&Holotype: &&&###Mexico.### \$\$\$Michoacán:\$\$\$ %%%Morelia,%% Hostel Allende, (E19.70181°, S101.19953°), P.H. Martins, 20–25/VII/2014, 1 #ff (CNA
5	<i>Filistatinella pistris</i>	&&&Non-type material: &&&###United States of America.### \$\$\$Arizona:\$\$\$ %%%Santa Rita Mts.,%% Madera Canyon, (E20.19°, S98.59°), A.M. & L.I. Davi
6	<i>Filistatinella tohono</i>	&&&Non-type material: &&&###United States of America.### \$\$\$Arizona:\$\$\$ %%%Peppersauce,%% [E37.08078°, S119.48541°], 3/VIII/1956, 1 #m (AMNH I

The tagged output of *Filistatinella hermosa* reads like this:

&&&Holotype: &&&###United States of America.### \$\$\$California:\$\$\$ %%%Los Angeles,%% Altadena, (E19.40954°, S99.72658°, 2618m), L. Pinter, 22/II/2010, 1 #m (MCZ 68568). &&&Paratypes: &&&###United States of America.### \$\$\$California:\$\$\$ %%%Santa Cruz Island,%% (E19.40954°, S99.72658°, 2618m), R.V. Chamberlin, 22/II/2010, 1 #ff 1 imm. (AMNH IFM-1232); %%%Santa Monica,%% (E19.58°, S98.51°), W. Ivie, 22/IV/1963, subadult female, 4 #ff (AMNH IFM-1147). &&&Non-type material: &&&###United States of America.### \$\$\$California:\$\$\$ %%%Hermosa Beach,%% [E16.10164°, S91.67694°], coll. illegible, 23/XII/1974, 1 #ff 5 imm. (AMNH IFM-1161); %%%Santa Monica,%% (E20.5°, S99.25°), W. Ivie, 22/IV/1963, 1 imm. (AMNH).

After applying formatting automatically using the tags, the final result is:

HOLOTYPE: UNITED STATES OF AMERICA. California: *Los Angeles*, Altadena, (E19.40954°, S99.72658°, 2618m), L. Pinter, 22/II/2010, 1 #m (MCZ 68568). **PARATYPES: UNITED STATES OF AMERICA. California:** *Santa Cruz Island*, (E19.40954°, S99.72658°, 2618m), R.V. Chamberlin, 22/II/2010, 1 #ff 1 imm. (AMNH IFM-1232); *Santa Monica*, (E19.58°, S98.51°), W. Ivie, 22/IV/1963, subadult female, 4 #ff (AMNH IFM-1147). **NON-TYPE MATERIAL: UNITED STATES OF AMERICA. California:** *Hermosa Beach*, [E16.10164°, S91.67694°], coll. illegible, 23/XII/1974, 1 #ff 5 imm. (AMNH IFM-1161); *Santa Monica*, (E20.5°, S99.25°), W. Ivie, 22/IV/1963, 1 imm. (AMNH).

The formula trims and corrects the most common mistakes (double spaces, consecutive commas, etc, but please check the contents and format of the output list carefully (and please do report bugs/doubts if you have any!)

Formatting the text using tags:

To use the tags to apply formatting, follow these steps (in MS Word 2007–2010):

- 1- Open the *Find and Replace* menu and click on the *More* button;
- 2- Check the *Use wildcards* option;
- 3- On the *Find what* field, indicate the tag you want to replace (e.g. \$\$\$*\$\$\$). The asterisk indicates that *any* text between dollar signs will be formatted.
- 4- Select the *Replace with* field but do not introduce any text in it. Instead, click the *Format* button and select which format you want to apply for that particular tag (e.g. bold, italics, small caps, etc.).
- 5- Click *Replace all*. The desired formatting should be applied throughout the manuscript.
- 6- Now we need to remove the tags. Select the *Replace with* field and click *No formatting*, leaving the field empty. Then, select the *Find what* field and introduce *only* the tag (e.g. \$\$\$). Click *Replace all* to remove that particular tag.
- 7- Repeat the process for each different tag you had applied to your text (e.g. \$\$\$, %%, ###, etc.).

