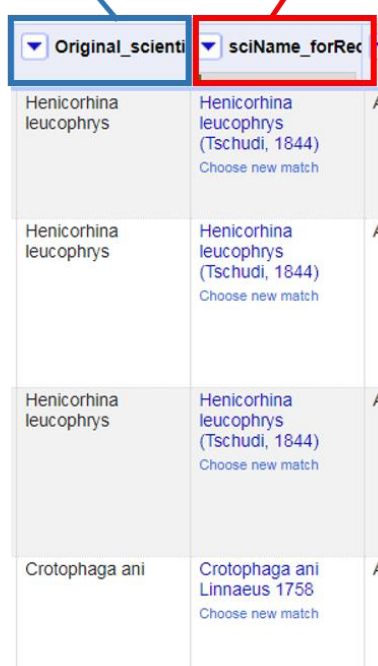


Algunos pasos posteriores a la reconciliación con EOL

1. Cómo retener los nombres científicos incluyendo autor

Luego de la reconciliación con EOL, tendremos una columna con los valores reconciliados:

scientificName original scientificName reconciliado



Original_scientificName	sciName_forRec
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match
Crotophaga ani	Crotophaga ani Linnaeus 1758 Choose new match

Por default, si bien vemos los autores, el valor de las celdas es sólo el nombre científico.

Para retener los autores, armaremos una nueva columna en base a la columna reconciliada:

Edit column → Add column based on this column...

Allí, nombramos la nueva columna y en el cuadro colocamos la siguiente expresión:

cell.recon.match.name

Esta expresión toma el nombre completo provisto por EOL, incluyendo autor.

Add column based on column sciName_forReconc

New column name: **scientificName_reconc_withAuthors**

On error: ☒ set to blank ☐ store error ☐ copy value from original column

Expression: **cell.recon.match.name** Language: Google Refine Expression Language (GREL) No syntax error.

Preview History Starred Help

row	value	cell.recon.match.name
146.	Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844)
247.	Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844)
260.	Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844)
353.	Crotophaga ani	Crotophaga ani Linnaeus 1758
414.	Colibri coruscans coruscans	Colibri coruscans coruscans (Gould, 1846)
448.	Colibri coruscans coruscans	Colibri coruscans coruscans (Gould, 1846)
481.	Colibri coruscans coruscans	Colibri coruscans coruscans (Gould, 1846)

OK Cancel

Tendremos entonces lo siguiente:

scientificName original	scientificName reconciliado	scientificName reconciliado incluyendo autor
Original_scientificName	sciName_forReconc	scientificName_reconc_withAuthors
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	Henicorhina leucophrys (Tschudi, 1844)
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	Henicorhina leucophrys (Tschudi, 1844)
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	Henicorhina leucophrys (Tschudi, 1844)
Crotophaga ani	Crotophaga ani Linnaeus 1758 Choose new match	Crotophaga ani Linnaeus 1758

2. Cómo capturar las URL que corresponden a cada nombre en EOL

Como vemos en el campo reconciliado, podemos acceder a las páginas correspondientes a cada nombre en EOL (haciendo click en los link). Si quisiéramos retener esas páginas de referencia, podemos armar un nuevo campo que contenga las URL correspondientes.

Para ello, a partir del campo reconciliado primero armaremos una nueva columna que retenga los identificadores de cada nombre en EOL (los IDs):

Edit column → Add column based on this column...

Allí nombramos la nueva columna y en el cuadro colocamos la siguiente expresión:

cell.recon.match.id

Add column based on column sciName_forReconc

New column name

EOL_ID

On error

☒ set to blank ☐ store error ☐ copy value from original column

Expression

Language Google Refine Expression Language (GREL) ▼

cell.recon.match.id

No syntax error.

Preview

History

Starred

Help

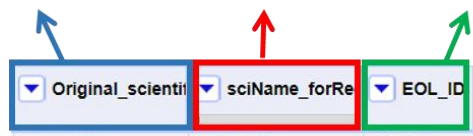
row	value	cell.recon.match.id
146.	Henicorhina leucophrys	1053020
247.	Henicorhina leucophrys	1053020
260.	Henicorhina leucophrys	1053020
353.	Crotophaga ani	915111
414.	Colibri coruscans coruscans	1227055
448.	Colibri coruscans coruscans	1227055
481.	Colibri coruscans coruscans	1227055

OK

Cancel

Tendremos entonces:

scientificName original scientificName reconciliado ID del nombre en EOL



Original_scientificName	sciName_forReconciliation	EOL_ID
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	1053020
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	1053020
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	1053020
Crotophaga ani	Crotophaga ani Linnaeus 1758 Choose new match	915111

A continuación, armamos una nueva columna en base a la columna de IDs:

Edit column → Add column based on this column...

Nombramos la nueva columna y en el cuadro colocamos la siguiente expresión:

"http://eol.org/pages/" + value

Esta expresión está armando una URL utilizando para cada celda el valor del ID correspondiente.

Add column based on column EOL_ID

New column name: **EOL_URL**

On error: ☒ set to blank ☐ store error ☐ copy value from original column

Expression: `"http://eol.org/pages/" + value` Language: Google Refine Expression Language (GREL) No syntax error.

Preview History Starred Help

row	value	"http://eol.org/pages/" + value
146.	1053020	http://eol.org/pages/1053020
247.	1053020	http://eol.org/pages/1053020
260.	1053020	http://eol.org/pages/1053020
353.	915111	http://eol.org/pages/915111
414.	1227055	http://eol.org/pages/1227055
448.	1227055	http://eol.org/pages/1227055
481.	1227055	http://eol.org/pages/1227055

OK Cancel

Tendremos entonces lo siguiente:

scientificName original	scientificName reconciliado	ID del nombre en EOL	URL para el nombre en EOL
Original_scientificName	sciName_forReconciliation	EOL_ID	EOL_URL
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	1053020	http://eol.org/pages/1053020
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	1053020	http://eol.org/pages/1053020
Henicorhina leucophrys	Henicorhina leucophrys (Tschudi, 1844) Choose new match	1053020	http://eol.org/pages/1053020
Crotophaga ani	Crotophaga ani Linnaeus 1758 Choose new match	915111	http://eol.org/pages/915111