Using `gnverify` on MS Windows 10

Dmitry Mozzherin

2020-04-21 16:39:27

Contents
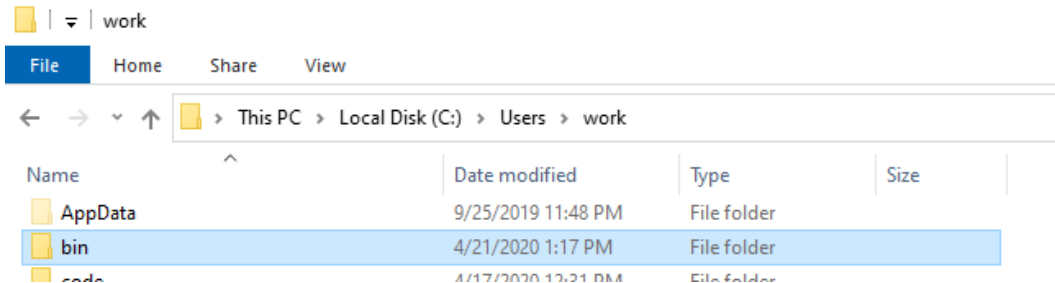
Installation

1. Go to `https://github.com/gnames/gnverify/releases/latest` and download zip file for windows (it will have a version higher than on the picture).
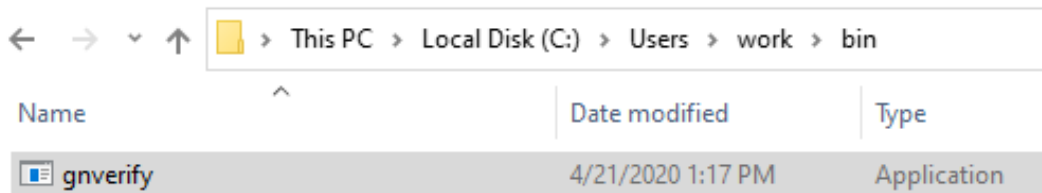
This zip file contains just one compressed executable file.

2. Create a folder called `bin` in your `home` folder (This is the folder with Downloads, Pictures, Desktop folders).

3. Unzip `gnverify.exe` file from the zip and place it into the `bin` folder.
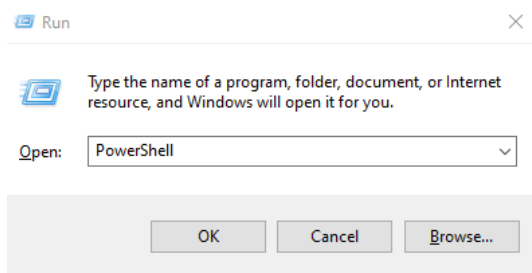

4. Create `data` folder in your home folder.

5. Open your data in Excel or Word and save scientific names, one per line or per row in a text file or tab-separated (TSV) file. Place this file into `data` folder. In our example, we will call the file `my-checklist.tsv`.

6. Note the position of the field that contains name-strings to verify. If you have it as the first/only field, no action is needed, but if it is the second, or the third etc. field in your tab-separated value, remember the number of the field. You will need to enter it for the **name_field** option.
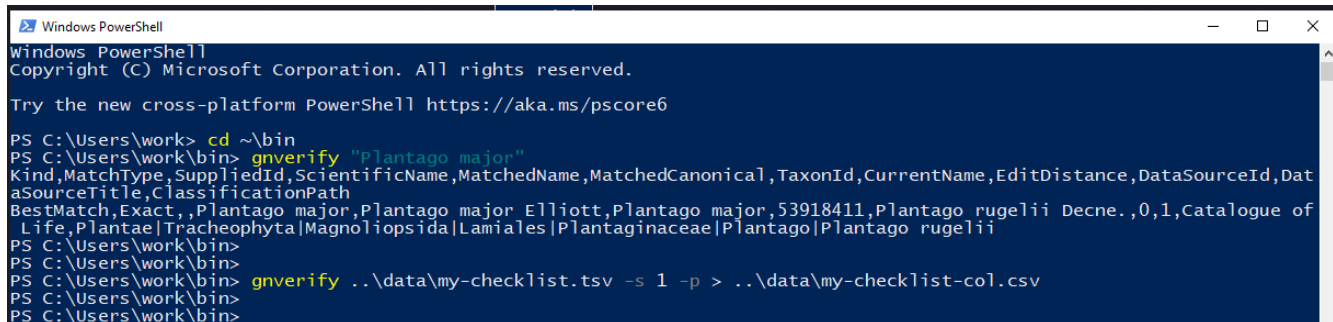
Usage

1. Push `Win-R` buttons on your keyboard to get a command prompt. Type there "PowerShell" and press the `OK` button.

2. You will see the PowerShell terminal window. Type the following commands:

```
cd ~\bin
gnverify "Plantago major"
```

After a short delay, you should see a result printed on the screen as a comma-separated value output. For the next step, we will read data from a file and save the resulting output into a new file.

3. To map data from your checklist to, for example, Catalogue of Life data, type:

```
gnverify ..\data\my-checklist.tsv -s 1 -p > ..\data\my-checklist-col-map.csv
```

This command will verify your names against data in Catalogue of Life and send resulting output to a new file. Now this file can be opened in Excel or Google Doc.

If the name-string field, often labelled as `ScientificName` is not the first/only field in your data file, you have to enter the number for that field using –name_field option. For example, if you have `ScientificName` as the third field in your file, enter:

```
gnverify ..\data\my-checklist.tsv -n 3 -s 1 -p > ..\data\my-checklist-col-map.csv
```

or

```
gnverify ..\data\my-checklist.tsv --name_field=3 --sources=1
        --preferred_only > ..\data\my-checklist-col-map.csv
```

4. To learn more about options of **gnverify** type

```
gnverify --help
```

```
Windows PowerShell                                                              —  □  >
PS C:\Users\work\bin> gnverify --help
gnverify 0.2.1
Verifies scientific names agains many sources.

USAGE:
    gnverify.exe [FLAGS] [OPTIONS] [INPUT]

FLAGS:
    -h, --help              Prints help information
    -p, --preferred_only    Ignores best match, returns only preferred results (if any).
    -V, --version           Prints version information

OPTIONS:
    -f, --format <format>      Format of the output: "compact", "pretty", "csv".
                               compact: compact JSON,
                               pretty: pretty JSON,
                               csv: CSV (DEFAULT)
    -s, --sources <sources>    IDs of important data-sources to verify against (ex "1,11").
                               If sources are set and there are matches to their data,
                               such matches are returned in "preferred_result" results.
                               To find IDs refer to "https://resolver.globalnames.org/resources".
                               1 - Catalogue of Life
                               3 - ITIS
                               4 - NCBI
                               9 - WoRMS
                               11 - GBIF
                               12 - Encyclopedia of Life
                               167 - IPNI
                               170 - Arctos
                               172 - PaleoBioDB
                               181 - IRMNG

ARGS:
    <INPUT>    A name-string or file to verify
```
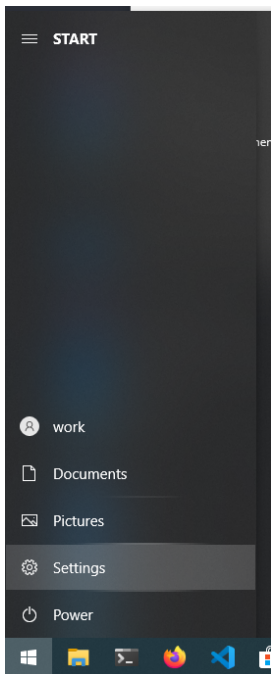
You can also read about gnverify at its github site
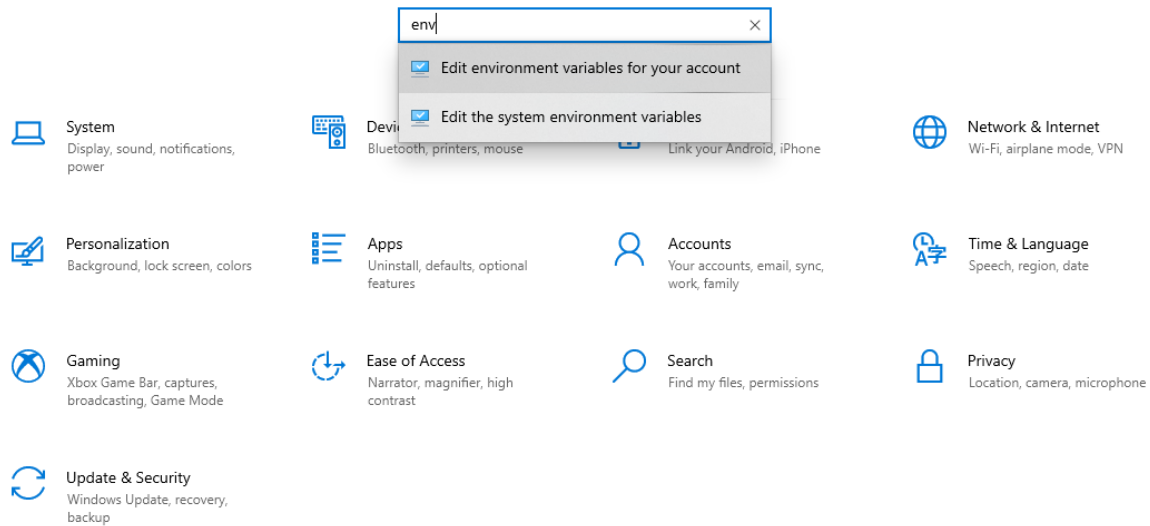
Making gnverify available from any folder

You can always run `gnverify` from the `bin` folder, but it is not very convenient. For example, you might like to run it from your data folder.

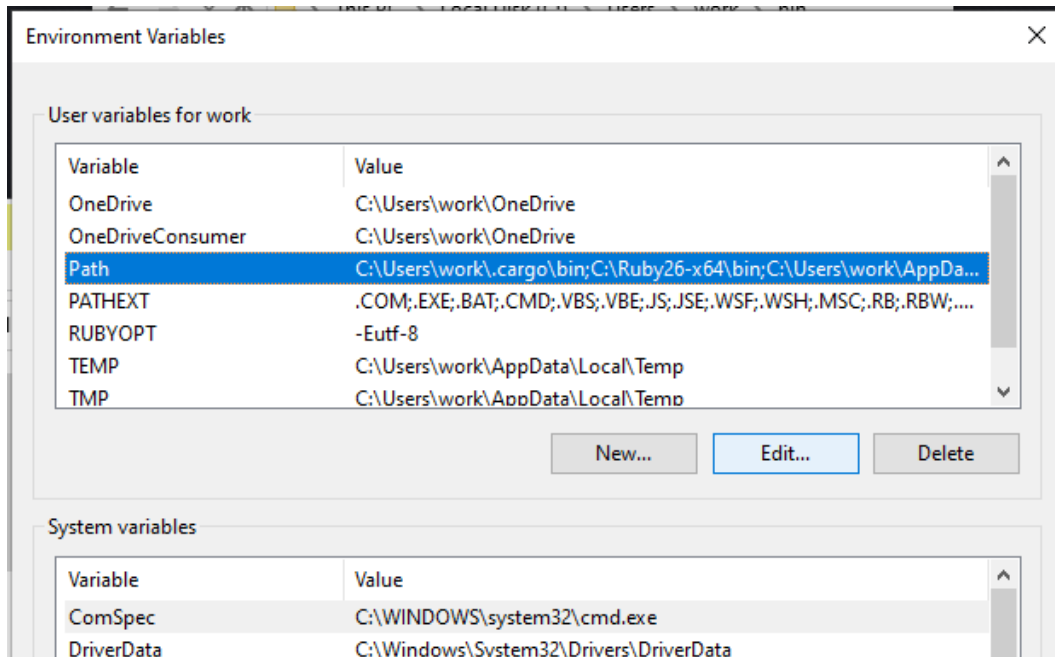1. To make `gnverify` available from any folder press the `Win` button and choose settings.

2. You will see settings windows with a search prompt in the center. Type "env" in the search prompt, and you will see a couple of returned results. Choose "Edit environment variables for your account".
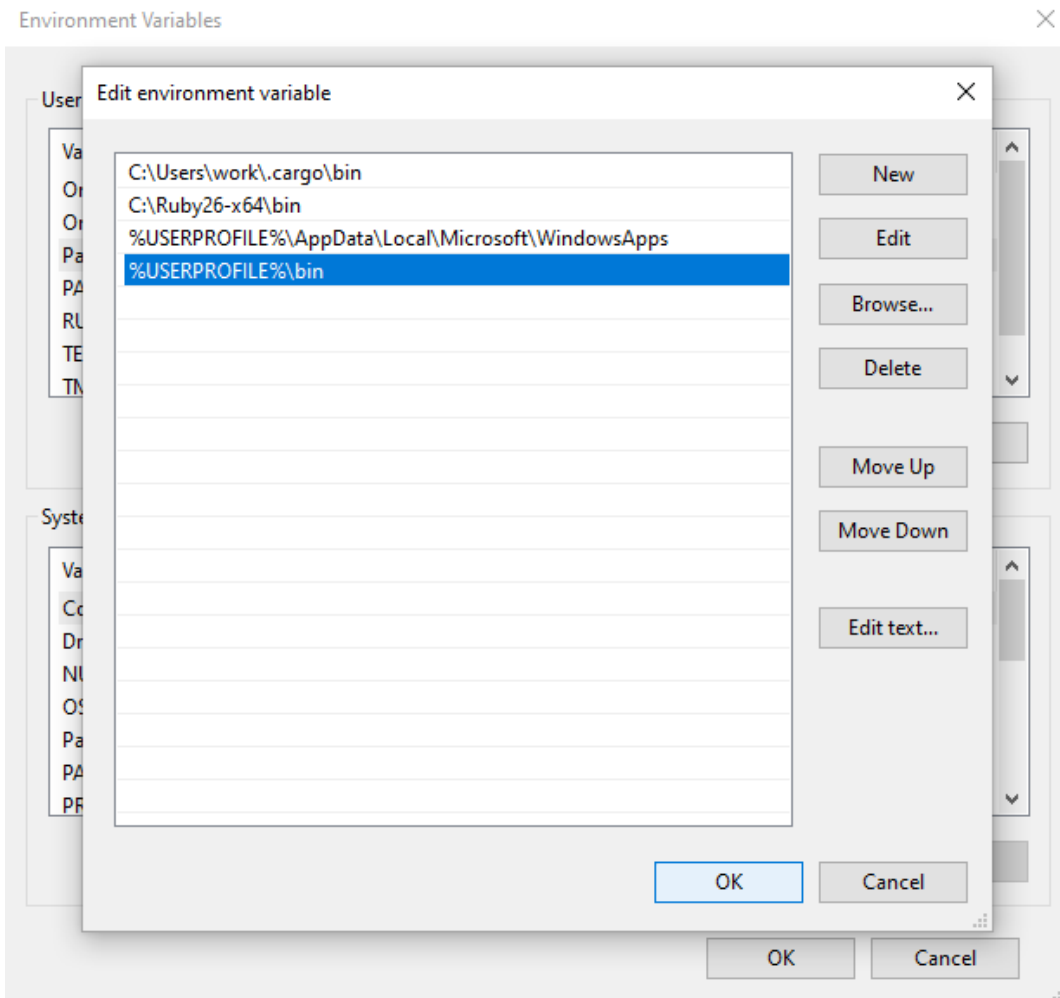
3. Now we need to find **Path** variable in the first window, select it and push the "Edit" button

4. When you see Edit window, push the "New" button and type in either a full path to the `bin` folder (Something like `C:\Users\my-user\bin`) or `%USERPROFILE%\bin`

Edit environment variable dialog:

- C:\Users\work\.cargo\bin
- C:\Ruby26-x64\bin
- %USERPROFILE%\AppData\Local\Microsoft\WindowsApps
- %USERPROFILE%\bin

Buttons: New, Edit, Browse..., Delete, Move Up, Move Down, Edit text..., OK, Cancel

5. Now you can use `gnverify` from any folder in `PowerShell` or `cmd` terminal window.

```
cd ~\data
gnverify my-checklist.tsv
```