
Subject: Issues with reading CSV-files

Posted by [Jan-Mathis](#) on Mon, 27 May 2019 08:54:09 GMT

[View Forum Message](#) <> [Reply to Message](#)

I downloaded a CSV-file from the ChEMBL website (see attachments). If I open it, DataWarrior falsely recognizes SMILES in the "ChEMBL ID" and "Type" column. Also the header is erroneous. For example it displays the title "Type" as "T_y_p_e".

I already downloaded the bugfix you posted on another thread on the 17.1.2019.

File Attachments

1)

[CHEMBL25-chembl_molecule-OjTXHbpRrTByJ-WgMV3ZWCbpTFbxBUjcW19lyizQMIM=.csv](#), downloaded 591 times

Subject: Re: Issues with reading CSV-files

Posted by [thomas](#) on Tue, 28 May 2019 08:48:39 GMT

[View Forum Message](#) <> [Reply to Message](#)

DataWarrior expects UTF-8 encoded text files. Your file is UTF-16 encoded. You may convert the file with many text editors, e.g. use "Text Editor" in Ubuntu to open the file. Then choose 'Save As', choose character encoding 'UTF-8' and press OK. I have attached the converted file, which opens nicely in DataWarrior.

File Attachments

1) [chembl25_problem_UTF8.csv](#), downloaded 511 times

Subject: Re: Issues with reading CSV-files

Posted by [Jan-Mathis](#) on Tue, 28 May 2019 19:13:29 GMT

[View Forum Message](#) <> [Reply to Message](#)

Thanks for the quick answer. Now everything is working.
