Subject: Re: Filter out nasty functions
Posted by thomas on Wed, 14 Jun 2023 19:30:03 GMT
View Forum Message <> Reply to Message

If Java strings contain a back slash '\' this needs to be encoded as double back slash '\\'. This explains, why three of the idcodes of the NastyFunctionDetector cannot be converted by DataWarrior. You need to replace '\\' again by '\' before pasting. You should also add a title line, e.g. "idcode<TAB>name". <TAB> should be a real TAB character and 'idcode' is a keyword for DataWarrior to tell it that this column contains idcode-encode chemical structures.

If you paste such a table with Edit->Paste, then DataWarrior creates a new document with the idcodes correctly decoded as substructure queries rather than complete structures. In this case the structure column is marked to contain substructures rather than structures. In this case it is important, because the nasty function substructures contain query features, which are lost if one copies the idcodes into an existing document's structure column. Then DataWarrior automatically converts substructures to structures, which removes atom/bond query features and considers open valences to be filled with hydrogen atoms.

I attach both, the text file with the idcodes and the dwar file containing the nasty function defining substructures.

Query feartures, which are not directly visible, can be recognized by their yellow atom/bond backgrounds. You can see all query features used by first opening the substructure in the structure editor and then opening the query feature dialog of a bond or atom. To open the editor double click the structure in the table. Then make sure the lasso tool is chosen in the editor. Then double click an atom or bond. The respective atom or bond query feature dialog opens and shows the current selection of query features for that atom or bond.

File Attachments
1) nasty_functions.txt, downloaded 372 times
2) nasty_functions.dwar, downloaded 502 times