Subject: Saving 2D-structure with labels
Posted by vijaymasand on Fri, 20 May 2022 15:04:06 GMT
View Forum Message <> Reply to Message

Dear all
Recently, I started using Datawarrior for research work, especially for QSAR and allied areas. It is really wonderful and useful for cheminformatics. I have three requirements from developers.
(1) I want to save the 2D-structure along with the labels, please see the ‘structure view’ panel in the attached image. I tried ‘copy view image’ but it copied only the structure and not the data labels. Is there any method to copy and paste (or save) the structure along with labels?
(2) Right now, it is providing data from Chembl 28. Recently, Chembl 30 was released. Is possible to force it to download data from Chembl 30?
(3) In ‘Look and Feel’ menu could you please add ‘white’ as an option.
Thanks in advance.
From
Vijay

File Attachments

1) Lables-datawarrior.png, downloaded 382 times

Subject: Re: Saving 2D-structure with labels
Posted by nbehrnd on Sun, 05 Jun 2022 14:35:25 GMT
View Forum Message <> Reply to Message

Dear Vijay,

addressing point 1) and a set of locally defined molecules to use the «copy new image» from a structure view, the export of structures and annotations as .png, or .svg worked well for me. Both exported illustrations contain a representation of the structure and the property previously selected:

The files exported were read easily by inkscape.[1] Here, I opted for the .png with a solid white background, though it is possible to opt-in for a transparent one. These observations refer to DW 5.5.0 for Linux in Debian 12/bookworm (branch testing) including DW's update by May 13, 2022. Thus, to ease replication of your problem, may you state the DW version you use and share a minimal .dwar file?

Norwid

[1] https://inkscape.org

File Attachments
1) example.dwar, downloaded 393 times
2) step.png, downloaded 1399 times
3) example_export.png, downloaded 368 times
4) example_export.svg, downloaded 414 times

Subject: Re: Saving 2D-structure with labels
Posted by vijaymasand on Mon, 06 Jun 2022 05:37:44 GMT
View Forum Message <> Reply to Message

Dear Norwid
Thank you for the reply. The solution provided by you is really simple and useful. I also developed images with required labels using your method without any problem. This also solved my third requirement of inclusion of ‘white’ as an option in 'Look and Feel'.
However, when I tried it for more than 100 molecules, I got error message that it is out of memory. So, it has some limitations. I am using Windows 11 (64 bit), 32 GB RAM, and Datawarrior 05.05.00.
For point number 2, I hope the developers will develop a solution to download data from ChEMBL30.
With regards
Vijay

Subject: Re: Saving 2D-structure with labels
Posted by nbehrnd on Mon, 06 Jun 2022 17:14:30 GMT
View Forum Message <> Reply to Message

Dear Vijay,

lacking access to Windows 11, I'm unable to replicate your findings. Thus, the operation system, or/and the executable of DW specifically released for Windows might contribute to this.

However, because DW* does not provide a progress bar when generating a .png or .svg from the array / structure view, a (perhaps hasty) perception may be anywhere between «the request to export the table wasn't yet launched» and «because of the computational load, the computer froze». In my case, it takes some time i) before an icon about the new file is displayed in the file browser, ii) the CPU monitor records a significant increase in computation, iii) the new file starts to be written, iv) writing all the data into the new file is completed. Do you consider a library of 5k random molecules DW generated a lager one (attached below)? The annotated structure array, exported as .svg is larger than 100 MB; for inkscape, this represents a larger file.

Difficult to tell if a sequential generation of these small illustrations (one frame with one molecular structure & selected description) as individual .png/.svg in a pattern like 001.png, 002.png, 003.png, etc. (or their .svg analogues) would be an advantage for you.

Norwid

*) tested with DW 5.5.0 for Linux including the update by 2022-06-04 in Debian 12/bookworm (branch testing)

File Attachments

1) setup_5k_molecules.png, downloaded 350 times
2) 5k_Random_Molecules.dwar, downloaded 381 times

Subject: Re: Saving 2D-structure with labels
Posted by vijaymasand on Tue, 07 Jun 2022 01:44:43 GMT
View Forum Message <> Reply to Message

Dear Norwid,
Once again thank you for solving another issue. I found that not selecting all molecules before using "Copy View Image", generates the error message. When I selected all molecules using Ctrl+A before using "Copy View Image", then it was very easy to save 2D-strcuture with labels for all molecules as a .svg or .png file.
However, I couldn't find a suitable method for sequential generation of multiple small illustrations (one frame with one molecular structure & labels) as individual .png/.svg in a pattern like 001.png, 002.png, 003.png, etc. (or their .svg analogues).
Again thank you for the support.
with regards
Vijay

Subject: Re: Saving 2D-structure with labels
Posted by vijaymasand on Tue, 07 Jun 2022 07:26:04 GMT
View Forum Message <> Reply to Message

Dear Norwid,
I found another simple way to export all 2D-structures with labels as a PDF file with a superb publication quality resolution.
First, I selected all the molecules in 'Structure view', then from 'File' menu option, I selected 'Print...'. This opened a new small window. In this new window, I opted for "Microsoft Print To PDF' and also checked 'Print to file'. Then, I was able to save it as pdf. However, it is mandatory to add file extension '.pdf' after the desired name for the file.
I hope this will be useful to other users.
Have a nice day.
Vijay

Subject: Re: Saving 2D-structure with labels
Posted by nbehrnd on Tue, 07 Jun 2022 17:30:25 GMT
View Forum Message <> Reply to Message

Dear Vijay,

indeed, the «print to pdf» is an attractive alternative. Not only that the depictions don't pixelate as in the .png bitmap, the .pdf retains a searchable text layer. This then allows to identify /the/ entry of particular interest by a label set (SMILES string, catalogue number, etc.) after the document was exported.

Norwid

Subject: Re: Saving 2D-structure with labels
Posted by thomas on Mon, 27 Jun 2022 16:45:19 GMT
View Forum Message <> Reply to Message

Dear Vijay,

the ChEMBL server was updated recently to ChEMBL version 30.

If you select a structure view, which shows labels, and if you then select print from the File menu and choose a PDF-printer, then you should get a nice paginated PDF document with as many pages as needed to cover all your structures. The following picture shows a fraction of the pdf...

File Attachments
1) t.png, downloaded 1331 times

Subject: Re: Saving 2D-structure with labels
Posted by vijaymasand on Wed, 29 Jun 2022 02:05:34 GMT
View Forum Message <> Reply to Message

Dear Thomas
Thank you for updating the server to ChEMBL30.
In my recently published paper (Pharmaceuticals 2022, 15, 745. https://doi.org/10.3390/ph15060745), I used molecular descriptors calculated by Datawarrior. The paper is full of mechanistic interpretations and new findings.
I am exploring Datawarrior. Thank you for creating Datawarrior.

Vijay

Subject: Re: Saving 2D-structure with labels
Posted by nbehrnd on Thu, 30 Jun 2022 18:03:12 GMT
View Forum Message <> Reply to Message

Dear Vijay,

congratulations to your publication.

Note, table 1 includes some examples of «counting patterns» which indeed exceed DW's set of pre-set statistics (e.g., counting small saturated carbon-only cycles) available via chemistry -> from chemical structure -> calculate properties. However, the development version of DW offers an additional functionality via chemistry -> from chemical structure -> add substructure count. Assisted by the program's sketcher, you may define countable pattern matching your needs like «excluding S of sulfone, the eighth atom should be an aromatic carbon atom»

By selection (lasso) and double click, the properties of the atoms in question may be specified further. The documentation calls this «substructure count»,[1] just prior to the section header «The Structure Filter».

With best regards,

Norwid

[1] https://openmolecules.org/help/chemistry.html#StructureFilte r

File Attachments

1) special_count_sulfone.png, downloaded 1139 times
2) special_count_sulfone.dwar, downloaded 358 times
3) atom_features.png, downloaded 318 times

Subject: Re: Saving 2D-structure with labels
Posted by vijaymasand on Sat, 02 Jul 2022 02:36:40 GMT
View Forum Message <> Reply to Message

Thank you, dear Norwid
I am aware about the 'substructure count' and 'calculate properties' functionalities, which are really excellent and useful. I will definitely use them in future work.
Have a nice day.
Vijay
