

cdk2024 #1: NWO Open Science grant for the Chemistry Development Kit

Egon Willighagen 

Published April 7, 2024

Citation

Willighagen, E. (2024). cdk2024 #1: NWO Open Science grant for the Chemistry Development Kit. In *chem-bla-ics*. chem-bla-ics. <https://doi.org/10.59350/ytkmr-0vv92>

Keywords

Grant, Cdk, Cdk2024

Abstract

We recently got awarded our second NWO Open Science grant (OSF23.2.097), this time for the Chemistry Development Kit (CDK). “We” here is me and Alyanne de Haan, René van der Ploeg, and Marc Teunis from Hogeschool Utrecht. The proposal has been submitted for public dissemination in RIO Journal, like we did with the first NWO Open Science grant.

Copyright

Copyright © Egon Willighagen 2024. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

chem-bla-ics

We recently got awarded our [second](#) NWO Open Science grant ([OSF23.2.097](#)), this time for the [Chemistry Development Kit](#) (CDK). “We” here is me and [Alyanne de Haan](#), René van der Ploeg, and [Marc Teunis](#) from Hogeschool Utrecht. The proposal has been submitted for public dissemination in [RIO Journal](#), like [we did](#) with the first NWO Open Science grant.

The project formally started on April 1 but we had our kick-off meeting in Maastricht on April 4-5. We were joined by Javier and on the second day by Marvin, and Ozan from our [BiGCaT research group](#) in Maastricht. During this hackathon, I gave a (repeat) [presentation](#) about the history of the CDK which also included the problem that software using the CDK does not always use the most recent version.

And that, upgrading tools using the CDK with the latest CDK version, is the main topic of this grant (work package 2, WP2). The full proposal has the focus list of tools, but most of it is also listed in [the issue tracker](#) we have set up as project management tool on GitHub.

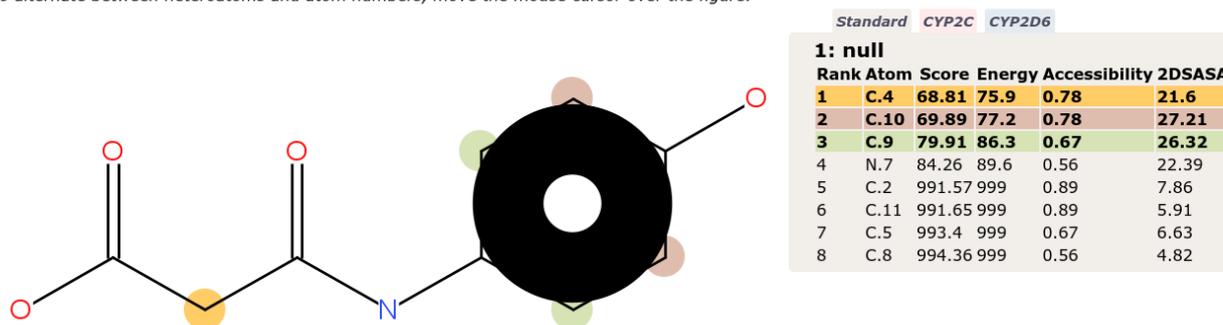
Second, we actually hacked together on two first tools, one on our focus list, but the other that was [requested we have a look at too](#): SMARTCyp. The latest version uses [RDKit](#) ([doi:10.1093/bioinformatics/btz037](#)), but the original version uses the CDK ([doi:10.1021/ml100016x](#)).

We downloaded the source code of SMARTCyp 2.4.2, started taking [notes](#), Javier [started](#) a Maven build environment, updated a lot of code, but we seem quite close to a version that can be tested by people that have integrated SMARTCyp in other tools. This is based on [CDK 2.9](#) and if you ignore the 2D depiction glitch, it looks it was a nice first choice:

Results from SMARTCyp version 2.5.0-SNAPSHOT

These results were produced: 2024-04-06_17-32-30. The infiles were: [smartcypxyz.smiles].

To alternate between heteroatoms and atom numbers, move the mouse cursor over the figure.



On a final note, we plan to record carefully our steps, in an open notebook science approach, with the intention to extract general upgrade steps. For example, we will update the [Migration](#) section of the [Groovy Cheminformatics with the Chemistry Development Kit](#).