chem-bla-ics

Groovy Cheminformatics 4th edition

Egon Willighagen

Published January 15, 2012

Citation
Willighagen, E. (n.d.). In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/yve3j-c9015

Keywords
Cdk, Cdkbook, Java, Cheminf

Copyright
Copyright © Egon Willighagen 2012. 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.

Six month was not quite the amount of time I anticipated between the third and fourth edition, but I finally managed to upload edition 1.4.7-0 of my Groovy Cheminformatics book. The first three editions sold 37 copies, including two for myself. Enough to feel supported and to continue working on it.

So, this new edition is again thicker, summing up to 152 pages now, which is 28 pages more than the 3rd edition . Indeed, the table of contents is more than half a page longer in itself, though, just barely, still fitting on four pages. In fact, I had to remove one (new) subsection title, because it would take otherwise two further pages.

The new content is again a mix of sections and chapters. While writing new chapters, I find myself realizing I need to cover more basics. Those get typically added as new sections. I did not get many feature requests, except for one email pointing me the text promised how to interpret and handle failing atom type perception, which explains one of the new sections. The full list of new content is:

- Section 2.1.4: explaining the three flavors of atomic coordinates
- Extended Section 2.2: added detail about electron counts of bonds (partly in reply to this post by Rich)
- Chapter 5 "Protein and DNA": four pages, mostly about PDB files, and the matching CDK data structure
- Chapter 6 "IChemObjectBuilders": four pages explaining the four alternative builders CDK 1.4.7 has
- Section 7.8: a new section with recipes on how to post-process read input, discussing MDL molfiles only now. It talks about what information is present in the file format, and what steps must be untertaken to add missing information
- Section 8.2.4 "No atom type perceived?!"
- Section 11.4: describes how to depict aromatic rings
- Section 11.5: describes how to change the background color of depictions
- Section 13.4: explains how to calculate the Van der Waals volume of molecules
- Section 18.1.3: discussing the API improvement in the iterating readers
- Appendix C: a list of all descriptors provided by the CDK
- Appendix D: a list of file formats known by the CDK, indicating which has readers and writers

On top of that, I improved other bits of the book too, such as the resolution of the depictions of molecules, as well as those of various diagrams. Also the number of scripts has seriously gone up, from 94 to 134!

Appendix C is a prelude to a chapter I am already writing, but did not get finished yet: a chapter about descriptor calculation. But since I just started a new post-doc position, it may take another six months for that chapter to make it into print.

The paperbak is available from Lulu.com, an on-demand publisher, as well as this ebook version.