chem-bla-ics

Legal Advice Needed: the NIH restricting access to our CC-licensed research results

Egon Willighagen

Published April 7, 2008

Citation
Willighagen, E. (n.d.). In *chem-bla-ics*. chem-bla-ics. https://doi.org/10.59350/865j7-crn03

Keywords
Pubmed, Pmc, Copyright

Copyright
Copyright © Egon Willighagen 2008. 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.

In reply to Peter's news that the NIH's PubMed Central (PMC) does not allow machine retrieval of content, I was wondering about this section in the CC license of much of the PMC content, such as our paper on userscripts (section 4a of the CC-BY 2.0):

> *You may not distribute, publicly display, publicly perform, or publicly digitally perform the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this License Agreement.*

CC-BY 3.0 reads differently, but has similar aims.

Let me make clear that I value machine readable publications much more than free (gratis, as-in-free-beer) publications. Now, the NIH initiative now just is 'Free Access'. An interesting step, but not one I care much about; not in relation to science anyway.

Now, Peter indicates that the NIH has put in place 'technological measures to control access' to the distribution of our work on userscripts (the PMC entry). That is in clear violation of the CC license.

I know that other NIH initiatives do allow this, such as PMC OAI, but that's just an 'auxiliary service'. It may come down to technical details, but some text on the PMC website is at least inaccurate:

> *Crawlers and other automated processes may NOT be used to systematically retrieve batches of articles from the PMC web site. Bulk downloading of articles from the main PMC web site, in any way, is prohibited because of copyright restrictions.*

They way it is described right now, it is like: *You may not drive a car*. Next paragraph. *But, if you have a driver license, we will approve*. Or, translated to this example: *You may only use this and that article, but only a few of them*. Next paragraph. *Unless you use the following technical hole in the measure we took to disallow you access.*

What the PMC website should indicate, instead, is that text mining is allowed for the PMC OAI subset, but that they would highly prefer to use the PMC OAI or PMC FTP routes. This is the least they have to do.

No matter what, I still have the feeling that any technical obstacles are disallowed by the CC-license. Any legal expert here, that can explain me if the CC license allows controlling how people have access to my material?