

## Packages - Freedom Issue #3308

### [cracklib] Proprietary wordlist

2022-07-13 03:54 PM - gap

<b>Status:</b> not-a-bug	<b>% Done:</b> 0%
<b>Priority:</b> bug	
<b>Assignee:</b>	
<b>Category:</b>	
<b>Description</b>	
This project's source repo includes a proprietary wordlist: <a href="https://github.com/cracklib/cracklib/tree/master/words">https://github.com/cracklib/cracklib/tree/master/words</a>	
Various other sources. Licensing of some of these lists is questionable - however I believe that these would reasonably fall under the 'list of facts' and 'can't copyright the phone book' copyright scenarios. If someone objects to a list being included here, it will be removed immediately.	
Copyright holder admits and attempts to rationalise his unjust monopoly:	
You're allowed to use and redistribute this wordlists collection or parts thereof, with or without modification, provided that credit is given where it is due, any modified versions are marked as such, this license is kept intact and included with each copy, and NO FEE IS CHARGED FOR OBTAINING A COPY except as negotiated with the copyright holder. In particular, you are NOT permitted to charge for bandwidth, physical media, and/or shipping. You're also not permitted to bundle this wordlists collection with a product you charge for.	
If redistribution for a fee is what you're after, please contact the copyright holder to negotiate special terms for the downloadable or the extended CD-ready version of this collection.	
It was a significant amount of work to compile this collection and having a monopoly on regulating the CD sales is my way to compensate for the time already spent and to allow for further work.	
-- Alexander Peslyak aka Solar Designer	
cracklib is required by 6 packages:	
<ul style="list-style-type: none"><li>- 389-ds-base</li><li>- deepin-pw-check</li><li>- lib32-cracklib</li><li>- libpwquality</li><li>- mariadb (optional)</li><li>- mariadb (make)</li></ul>	

### History

#### #1 - 2022-07-13 08:29 PM - bill-auger

On Wed, 13 Jul 2022 15:54:56 +0000 [labs@parabola.nu](mailto:labs@parabola.nu) wrote:

You're also not permitted to bundle this wordlists collection with a product you charge for.

that is definitely non-free

In particular, you are NOT permitted to charge for bandwidth, physical media, and/or shipping.

that is more like nonsense - none of those activities are related to copyright; so the authors can not prohibit those activities

but that license is for only the two files under words/files/openwall/ - the package itself could probably be rescued - unfortunately, that README give no licensing information for any of the other word lists - the files under words/files/cracklib/ are presumably under the cracklib license; but the others would need to be investigated

words/files/clarkson.edu/

- clarkson.txt.bz2

words/files/mark-burnett/

- 10k-most-common.txt.bz2
- passed-basic-check.bz2
- scored-above-50.bz2

words/files/skullsecurity.org/

- 500-worst-passwords.txt.bz2
- alypaa.txt.bz2
- cain.txt.bz2
- carders.cc.txt.bz2
- conficker.txt.bz2
- elitehacker.txt.bz2
- facebook-pastebay.txt.bz2
- facebook-phished.txt.bz2
- faithwriters.txt.bz2
- hak5.txt.bz2
- hotmail.txt.bz2
- myspace.txt.bz2
- phpbb.txt.bz2
- porn-unknown.txt.bz2
- singles.org.txt.bz2
- tuscl.txt.bz2
- twitter-banned.txt.bz2

### #2 - 2022-07-13 09:30 PM - gap

If such wordlists can indeed be copyrighted, I'd recommend immediately blacklisting cracklib whilst we work on a cleaned-up version, and report the issue upstream.

Unfortunately, from a quick check, it looks like blacklisting cracklib will have a catastrophic knock-on effect which might render Deepin and GNOME uninstalleable until the cleaned-up -libre version is complete.

### #3 - 2022-07-13 10:09 PM - bill-auger

Issue [#3308](#) has been updated by gap.

I'd recommend immediately blacklisting cracklib whilst we work on a cleaned-up version

that is not the blacklisting procedure though

- 1) evaluate and document the extent of the conflict
- 2) try to liberate the package

if that is not feasible:

- 3) try to find a suitable replacement

if that is not feasible:

- 4) blacklist it, as a last resort, because that may entail blacklisting other packages. each of which would need to go through this entire procedure, recursively

and report the issue upstream.

that is always a good idea after step 1 is completed (successfully or not) - if any of those other word lists are also non-free, then it would not solve the problem to report only the obvious ones

in this case, i think that is unlikely to be fruitful though; because the cracklib maintainers are already aware of the conflict, and have disclaimed the conflict explicitly - only their contributors or the authors of the non-free files could make them change their position

arch could do something about it though - if those files are essential to the program's operation, then the GPL is invalid

### #4 - 2022-07-13 10:56 PM - gap

Whilst that procedure is certainly technically superior, I thought our first priority was to comply with the GNU FSDG even if that means temporarily breaking the repos.

**#5 - 2022-07-13 11:22 PM - bill-auger**

this is the FSDG requirement:

So we expect distros to occasionally contain mistakes: nonfree software that slipped through  
Our requirement is for the distribution developers to have a firm commitment to promptly correct any mistakes that are reported to them.

the "mistake" was only discovered today; so we are still well within the range of "promptly"

if you have the time, download all those word lists and check if they have license files inside - if none of them do, then this package probably can not be rescued

**#6 - 2022-07-14 12:25 AM - gap**

I think we need FSF assistance as to whether such wordlists are even eligible for copyright.

**#7 - 2022-08-27 05:34 PM - GNUtoo**

Maybe if we have our own arrangement of words and add words to the list we could be OK?

Or maybe newer wordlists could be found in other packages like packages for cracking passwords like john the ripper. Using plain dictionaries would probably be too limited.

Denis.

**#8 - 2022-08-27 05:35 PM - GNUtoo**

Debian also has to only ship free software so they might also have done things to fix the cracklib package.

**#9 - 2022-08-27 05:41 PM - GNUtoo**

Ah I was mistaken in my interpretation: it seems that we are fine.

As I understand cracklib says that the password list is fine because

these would reasonably fall under the 'list of facts' and 'can't copyright the phone book' copyright scenarios

If I understood right, the restrictions only comes from the "Openwall word lists" and since cracklib has words from that list but also has words from different list as well, we should be safe.

Reference: <https://raw.githubusercontent.com/cracklib/cracklib/master/words/README.md>

**#10 - 2022-08-27 05:43 PM - GNUtoo**

- *Status changed from unconfirmed to not-a-bug*

Feel free to reopen this bug if I was wrong about the interpretation of the wordlist license or if I missed facts.

**#11 - 2022-08-27 07:08 PM - gap**

The issue I see is that the people who put these word lists together are claiming copyright and putting restrictions on them, so if we dispute this on the basis that they are not copyrightable in the first place, we probably need to get in contact with those people and raise the issue, instead of putting Parabola users at risk who might be sued by them.

Again, I think we need FSF assistance.

**#12 - 2022-08-27 10:26 PM - bill-auger**

Issue [#3308](#) has been updated by GNUtoo.

As I understand cracklib says that the password list is fine because

these would reasonably fall under the 'list of facts' and 'can't copyright the phone book' copyright scenarios

ah but; the FSF says that you can copyright a phone book, and people do - this grey area is probably only going to get grey-er, as data-mining is becoming such a big business - companies already want to copyright their collected data sets, regardless of how trivial or factual or easily captured each bit of data is

**#13 - 2022-09-01 06:29 PM - gap**

I don't think closing this issue is a solution.

I agree with [bill-auger](#), and his previous message would imply we should try to liberate it or find a replacement, as he said before, but since nobody appears to want to do this at the moment, I think we should blacklist it instead of leave it lingering.

These sorts of grey areas and the resulting indecision seems to be taking a toll on our psyche like in [#1035](#), so I'd suggest we make the decision to blacklist when in doubt, as making a reversible decision to blacklist so the issue is at least temporarily dealt with is a better solution than leaving an issue lingering, in my opinion.

Indeed, the blacklist contains many packages that were blacklisted without the procedure of attempting liberation and then attempting to find a replacement was done.

**#14 - 2022-09-02 03:19 AM - bill-auger**

any adequate replacement would necessarily have the same list of common passwords - for that reason, I really don't think those password lists are copyrightable

I would not argue about removing it though, simply because it is an ugly tool anyways - its only valid use-case is cracking someone else's password - calamares uses it to warn the user that the chosen password is "weak" or "too common"; but that is a dumb use-case

calamares originally refused to install the OS unless the user selected a password which is not on the cracklib crack list - so, I added a new checkbox to select "but that's ok - I really *do* want a weak password - please don't prevent me from doing what I want"