OPEN SOURCE COMPLIANCE

OSADL stellt FOSS-Kuratierungsdatenbank vor

OSADL eG stellt eine frei verfügbare Kuratierungsdatenbank vor. Wofür braucht man diese und wie viel Arbeit kann eingespart werden, wenn man bestehende Kuratierungsdaten verwendet?

Die zunehmende Verbreitung und Akzeptanz von Software-Komponenten, die unter einer sogenannten Open-Source-Lizenz stehen (Free and Open Source Software, FOSS), wird allgemein darauf zurückgeführt, dass diese Lizenzen die uneingeschränkte Wiederverwendung und Verbreitung der Software ermöglichen – und dies sogar nach deren Veränderung oder Erweiterung. Um diese Rechte nutzen zu können, sind zwar keine Lizenzgebühren zu entrichten, aber es sind Lizenzpflichten dahingehend zu erfüllen, dass dem Empfänger der Software Rechte eingeräumt und dafür bestimmte Dokumente verfügbar gemacht werden.

Da ein Anwender von FOSS normalerweise nicht in direktem Austausch mit den jeweiligen Rechteinhabern steht, können die Lizenzbedingungen nicht individuell ausgehandelt werden, sondern müssen so erfüllt werden, wie sie sind. Dies hat aber den Vorteil, dass alle Anwender die gleichen Pflichten erfüllen müssen. Dennoch leisten alle Unternehmen, die FOSS in ihren Produkten einsetzen, üblicherweise jeweils allein die gleiche Dokumentations-

Projekt OSSelot: Der Ozelot gab der FOSS-Kuratierungsdatenbank von OSADL ihren Namen.

Bild: Leonardo - stock.adobe.com

arbeit, um ein Produkt mit FOSS für den Vertrieb freizugeben. Diese unnötige Parallelarbeit führt dazu, dass ein Teil der ökonomischen Vorteile von FOSS wieder verlorengeht.

Es lag also nahe, ein Projekt ins Leben zu rufen mit dem Ziel, eine frei verfügbare Datenbank zu entwickeln, die alle Dokumente enthält, die für die lizenzkonforme Weitergabe häufig verwendeter FOSS-Pakete erforderlich sind. Damit dann in der praktischen Arbeit auch tatsächlich auf diese kuratierten Daten zurückgegriffen wird, muss die Herstellung nach allgemein akzeptierten Kriterien erfolgen und diese müssen so weitgehend transparent gemacht werden, dass das nötige Vertrauen in die Daten entsteht.

Bild: OSADL eG

VERFASST VON
Caren Kresse
Compliance und Technologie
OSADL eG

Bereitstellung von Kuratierungsdaten

Ein unverzichtbarer Schritt in jedem FOSS-Compliance-Prozess ist die Ermittlung der geltenden Lizenzen für die verwendeten Softwarekomponenten. Zu diesem Zweck werden die Lizenzinformationen aus dem Quellcode extrahiert und aufgelistet. Diese Informationen können dann im Clearing-Prozess verwendet werden, d. h. bei der Entscheidung, ob die Lizenzverpflichtungen der Software für den beabsichtigten Anwendungsfall erfüllt werden können. Darüber hinaus verlangen fast alle FOSS-Lizenzen, dass Lizenztext und Copyright-Hinweise mit der Software mitgeliefert werden.

Wenn die Software in binärer Form geliefert wird, müssen diese Materialien separat bereitgestellt werden. Außerdem erfordern manche FOSS-Lizenzen in diesem Fall, dass der komplette korrespondierende Quellcode sowie Build- und Installations-Skripte verfügbar gemacht werden. Mit Ausnahme der Installations-Skripte können alle diese Materialien in allgemeiner Form hergestellt und von Anwendern wiederverwendet werden, ohne dass individuelle Anpassungen erforderlich sind.

Haftungsrisiko und Vertrauen

Wenn Compliance-Materialien öffentlich zugänglich gemacht und von Unternehmen für ihre eigenen Produkte wiederverwendet werden, dann muss das Haftungsrisiko so gering sein, dass es gerechtfertigt ist, diesen Daten zu vertrauen. Haftungsrisiko und Vertrauen sind also zwei Seiten derselben Medaille; denn nur wenn erkennbar ist, dass die Daten sorgfältig aufgearbeitet – kuratiert – wurden, wird das Haftungsrisiko ausreichend minimiert.

Dazu gehört, dass die Kuratierung von namentlich genannten Personen mit entsprechendem Fachwissen durchgeführt wird und dass es einen transparenten Review-Prozess gibt. Dabei darf nicht vergessen werden, dass die Personen, welche die Datenkuratierung durchführen, mit ihrem Namen dafür geradestehen, dass sie mit größtmöglicher Sorgfalt gearbeitet haben. Und nicht zuletzt bürgt auch die Organisation, welche die Materialien bereitstellt, mit ihrem Ruf für die Qualität des gesamten Projekts.

Diese Personen und Organisationen sind daher dem Risiko ausgesetzt, in einem Rechtsstreit wegen Urheberrechtsverletzungen, die sich aus falschen Informationen in bereitgestellten Compliance-Materialien ergeben können, zu haften.

Dieses Haftungsproblem kann dadurch gemildert werden, dass die Materialien unter einer Lizenz zur Verfügung gestellt werden, die bedingungslos maximale Rechte ge-

```
FileName: coreutils-9.1/coreutils-9.1/coreutils-9.1/src/blake2/blake2b-ref.c
FileChecksum: SHA1: 63ed5dcfd161a4fa33d9415c85bb753e9f8b7e42
FileChecksum: SHA256: 325956565a15d4a1cf91beee2676c8bef0316a7a3d2b95c3e71a1e5d87ae0d47
FileChecksum: MD5: 28d09caefc730ba5ec0593f4ff5ef776
LicenseConcluded: Apache-2.0 OR CC0-1.0 OR OpenSSL
LicenseComments: <text>The information in the file is:

You may use this under the
terms of the CC0, the OpenSSL Licence, or the Apache Public License 2.0, at
your option. The terms of these licenses can be found at:

- CC0 1.0 Universal : https://creativecommons.org/publicdomain/zero/1.0
- OpenSSL license : https://www.openssl.org/source/license.html
- Apache 2.0 : https://www.apache.org/licenses/LICENSE-2.0 </text>

Thus it is a "dual licensing" case. We have chosen CC0-1.0 due to
compatibility with GPL-3.0.
Remark: The OpenSSL license mentioned here is the "old" OpenSSL license,
the copyright statement in the file is dated 2012. In this time the only
valid license that had this name was the OpenSSL license (if you follow
the link today you will see that from version 3.0 onward OpenSSL is
licensed under Apache-2.0).
The link was visited on 21st of Sept 2022.
LicenseInfoInFile: Apache-2.0
LicenseInfoInFile: OpenSSL
LicenseInfoInFile: CC0-1.0
FileCopyrightText: <text> Copyright 2012, Samuel Neves <sneves@dei.uc.pt> </text>
```

Bild: OSADL eG

SPDX-Beispiel: Kuratierte Compliance-Informationen, die eine SPDX Tag:Value-Datei für eine Quellcodedatei enthält.

währt. In diesem Fall gilt das Schenkungsrecht, wonach Haftung nur bei Vorsatz und grober Fahrlässigkeit greift. Die Creative Commons Zero 1.0 Universal (CC0-1.0) ist ein Beispiel für eine solche Lizenz.

Die Sorge um Haftungsfälle in Bezug auf öffentlich zur Verfügung gestellte, rechtlich relevante Informationen im FOSS-Umfeld war bis vor einigen Jahren vor allem in den Vereinigten Staaten sehr groß. Da jedoch diesbezüglich keine relevanten Gerichtsverfahren stattfanden oder bekannt wurden, sind die Vorbehalte deutlich geringer geworden. In Verbindung mit der oben beschriebenen großzügigen Lizenzierung ist es daher inzwischen denkbar, dass sich eine Community zur öffentlichen Bereitstellung und Nutzung von Compliance-Materialien bildet.

Darstellung von Compliance-Materialien

Damit Compliance-Materialien tatsächlich von verschiedenen Nutzern wiederverwendet werden können, muss eine möglichst universelle Darstellungsform gewählt werden, die idealerweise auch leicht konvertierbar ist. Als Grundlage bietet sich hierfür das einfache Textformat an. Dieses kann dann nach einem etablierten Standard, zum Beispiel dem Software Package Data Exchange (SPDX) Tag:Value-Standard, formatiert sein.

Hierbei werden für jede einzelne Quellcodedatei eines Softwarepakets der Pfad der Datei (mit dem Tag „FileName“), diverse Checksummen („FileChecksum“), die von den Scanning-Tools gefundenen Lizenzen („LicenseInfoInFile“), die letztendlich durch die Kuratierung bestimmten Lizenzen („LicenseConcluded“), gegebenenfalls ein Kommentar zur Begründung einer Lizenzentscheidung („LicenseComments“) und die Urhebervermerke („FileCopyrightText“) aufgelistet. Die Lizenzen werden mit einem Kürzel bezeichnet, der dazugehörige Lizenztext wird dann einmalig am Ende der Datei aufgeführt.

Ein Beispiel für einen solchen Eintrag zeigt das Bild. Die Informationen in der genannten SPDX Tag:Value-Datei sind sowohl für Menschen lesbar als auch für die automatisierte Verarbeitung geeignet. Zum Beispiel könnte ein Build-Prozess so erweitert werden, dass für jede Datei, die für ein bestimmtes Binärprodukt verwendet wird, über die Checksumme der Datei die Compliance-Informationen aus der SPDX Tag:Value-Datei herausgefiltert werden. Somit erhält man eine Zusammenstellung aller tatsächlich für das konkrete Binärprodukt anwendbaren Lizenzen.

Darüber hinaus kann die SPDX Tag:Value-Datei (gegebenenfalls mit Zwischenschritten zur Konvertierung) in eigene Instanzen verschiedener Tools importiert werden, um die Kuratierung anderer Versionen derselben Software zu vereinfachen und zu beschleunigen; denn in diesem Fall müssen nur noch die geänderten Dateien bewertet werden. Neben der maschinell verwendbaren SPDX Tag:Value-Datei, welche die Informationen für jede einzelne Datei enthält, können die Informationen, die zur Erfüllung der Lizenzpflichten mitgeliefert werden müssen, auch pro Komponente gebündelt werden. Dieses als „Disclosure-Dokument“ bezeichnete Format kann dann unmittelbar mit einem Produkt mitgeliefert werden, wie es FOSS-Lizenzen erfordern.

Wie viel Arbeit kann eingespart werden?

Wenn für eine Software-Komponente in einer bestimmten Version die Kuratierungsdaten bereits komplett vorliegen, ist die Aussage einfach, wie viel Arbeit eingespart werden kann – nämlich die gesamte Arbeit. Interessanter ist es, einmal abzuschätzen, wie viel Aufwand durch die Wiederverwendung von kuratierten Compliance-Materialien eingespart werden kann, wenn nur die Daten einer anderen Version vorliegen.

Dies wurde am Beispiel der OpenSSL-Bibliothek untersucht, und zwar wurde Mithilfe der SPDX Tag:Value-Datei der bereits kuratierten Version 3.0.5 die Version 3.0.7 neu kuratiert. Es zeigte sich, dass 227 von insgesamt 3329 Dateien neu analysiert werden mussten, also die Daten von 93 Prozent der Dateien wiederverwendet werden konnten. In der gleichen Größenordnung dürfte auch die Einsparung an Arbeit und Zeit liegen.

OSADL und das Projekt „OSSelot“

Das Open Source Automation Development Lab (OSADL) hat es sich zur Aufgabe gemacht, Materialien und Dienstleistungen bereitzustellen, die erforderlich sind, um Open Source-Software in industriellen Produkten einzusetzen. Dabei liegt der Schwerpunkt auf Leistungen, die nur einmal entwickelt, dann aber von vielen unverändert genutzt werden können. Insofern war es naheliegend, dass das beschriebene Projekt einer FOSS-Kuratierungsdatenbank vom OSADL eingebracht und organisiert wird.

Der Name des OSADL-Projekts lautet „OSSelot“. Bereits jetzt bietet dieses Community-Projekt Compliance-Materialien für 120 verschiedene Softwarepakete. Weitere Pakete werden regelmäßig hinzugefügt, und es wird erwartet, dass bis Mitte 2023 ein komplettes GNU/Linux-basiertes Embedded-System ausschließlich mit der Kuratierungsdatenbank lizenziert werden kann.

Die Compliance-Materialien sind auf GitHub unter https://github.com/Open-Source-Compliance/package-analysis öffentlich zugänglich und können unter der genannten Lizenz CC0-1.0 verwendet werden. Zusätzliche Informationen, Werkzeuge und Dokumentation sind auf der Projektseite von OSADL unter https://www.osselot.org verfügbar. *(jw)*