

Code aus dem Internet in der kommerziellen Softwareentwicklung

Recycling in der Firma

Eine wissenschaftlich fundierte Umfrage unter hauptberuflichen Softwareentwicklern hat recht unterschiedliche Einstellungen zum Basar-artigen Wiederverwenden von frei verfügbarem Quellcode ergeben. Erfahrungen, Rollen in der Firma und Programmiersprachen bestimmen die Akzeptanz. *Manuel Sojer*



Das Internet ist voll von Quellcode. Laut der Datenbank von Black Duck Software [1] gibt es heute mehr als 600 000 veröffentlichte Open-Source-Projekte, die zusammen 100 Milliarden Zeilen Quellcode und damit das Äquivalent von 100 Millionen Mannjahren Entwicklungsarbeit geschaffen haben. Hinzu kommen verfügbare Komponenten und Snippets, die nicht unter freien Lizenzen stehen, aber unentgeltlich nutzbar sind.

Seit den 60er Jahren gilt Wiederverwendung (Reusing) als zentral, um Software von hoher Qualität schnell und wirtschaftlich zu schaffen [2]. Solcher Code kam entweder aus dem eigenen Unternehmen oder von Dritten [3]. Beide Ansätze führen zu direkten Kosten, entweder für die Entwicklung In-House, die sich durch hoffentlich mehrfaches Verwenden amortisiert, oder für den Kauf von Komponenten.

Allein wegen der Kosten bildet unentgeltlich wiederverwendbarer Code aus dem Internet eine echte Alternative. Zudem entfallen bei Code aus gut geführten Open-Source-Projekten die Kosten für die zukünftige Codepflege.

Attraktiver Fundus

Darum überrascht es wenig, dass Code aus dem Internet für professionelle Software-Entwickler in Firmen – wie auch für Hobbyprogrammierer und Open-Source-Entwickler [4] – einen attraktiven Fundus bildet. Wie bei externem Code jedes Ursprungs stellt sich auch bei heruntergeladenem Code die Frage nach Qualität und Sicherheit. Obwohl es noch keine finalen Erkenntnisse gibt, legen neuere Untersuchungen [5],[6] nahe, dass es bei wohlüberlegter Wahl der Quellen keine grundsätzlichen Unterschiede zwischen kommerziell erworbenem und unentgeltlichem Code gibt.

Ein im Herbst 2009 unter professionellen Softwareentwicklern durchgeführte Umfrage sollte die praktische Seite des Reusing sowie die Handhabung möglicher rechtlicher Risiken untersuchen. Um eine möglichst heterogene Grundgesamtheit bestehend aus Softwareentwicklern verschiedener Länder, Unternehmen und Entwicklungsrollen zu erreichen, wurden Teilnehmer aus Newsgroups wie [[http://comp.lang.c++](http://comp.lang.c++.)], [<http://comp.lang.java.programmer>] und [<http://comp.lang.ruby>] zufällig ausgewählt. Die Implementierung der Online-Umfrage passierte mit Hilfe von Lime Survey und mit hauptsächlich geschlossenen, über Likert-Skalen zu beantwortenden Fragen.

Insgesamt nahmen an der Umfrage 869 professionelle Softwareentwickler teil. 84 Prozent waren zu der Zeit in diesem Beruf aktiv, der Rest hatte ihn früher ausgeübt. Mit der Hälfte und einem Viertel stellten Europa und Nordamerika die Hauptherkunftsregionen der zu 98 Prozent männlichen Teilnehmer dar, die im Durchschnitt 36 Jahre alt waren. Die Softwareentwickler blickten im Schnitt auf zehn Jahre Berufserfahrung zurück. 56 Prozent waren nebenher als Softwareentwickler in einem oder mehreren öffentlichen Open-Source-Projekten involviert. Für 20 Prozent der Teilnehmer stellte C++ die primäre Programmiersprache dar, für 14 Prozent Java und für 12 Prozent Python. Weitere Informationen zur Umfragetechnik und Demographie sind in [7] veröffentlicht.

Wachsende Bedeutung

Auf einer fünfstufigen Skala von "Überhaupt nicht wichtig" bis "Sehr wichtig" schätzten die befragten aktiven Entwickler das Thema Wiederverwendung durchschnittlich als "ein wenig wichtig" für ihre Arbeit ein (siehe [Abbildung 1](#)), nur 12 Prozent bezeichneten es als "überhaupt nicht wichtig" während für 19 Prozent Reusing sehr "sehr wichtig" war.

Fügt man der Betrachtung durch Berücksichtigen der ehemaligen Entwickler eine zeitliche Dimension hinzu, ist erkennbar, dass das Thema über die Jahre bedeutender wurde (siehe [Abbildung 2](#)). Umfrageteilnehmer, die bereits vor 2004 aufgehört hatten Software zu entwickeln, bewerteten Open Source meist zwischen "überhaupt nicht wichtig" und "kaum wichtig". Erst nach 2003 begann die Wichtigkeit, bis sie schließlich 2008 und 2009 eine durchschnittliche Relevanz von "Ein wenig wichtig" erreichte. Dies fällt zeitlich mit der Verbreitung von Open-Source-Software zusammen.

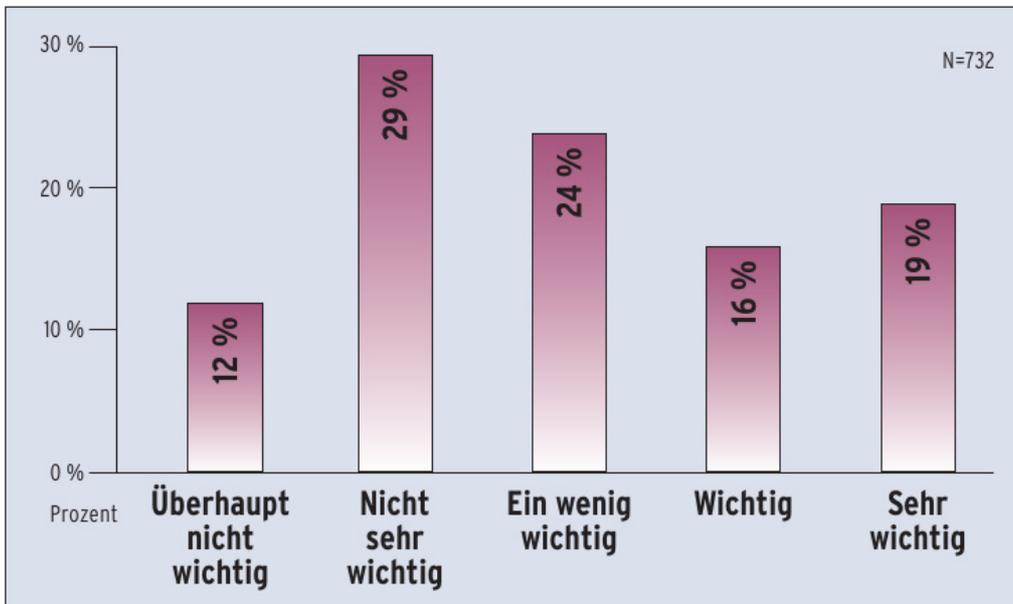


Abbildung 1: Welche Bedeutung messen professionelle Software-Entwickler der Wiederverwendung von Code aus dem Internet bei?

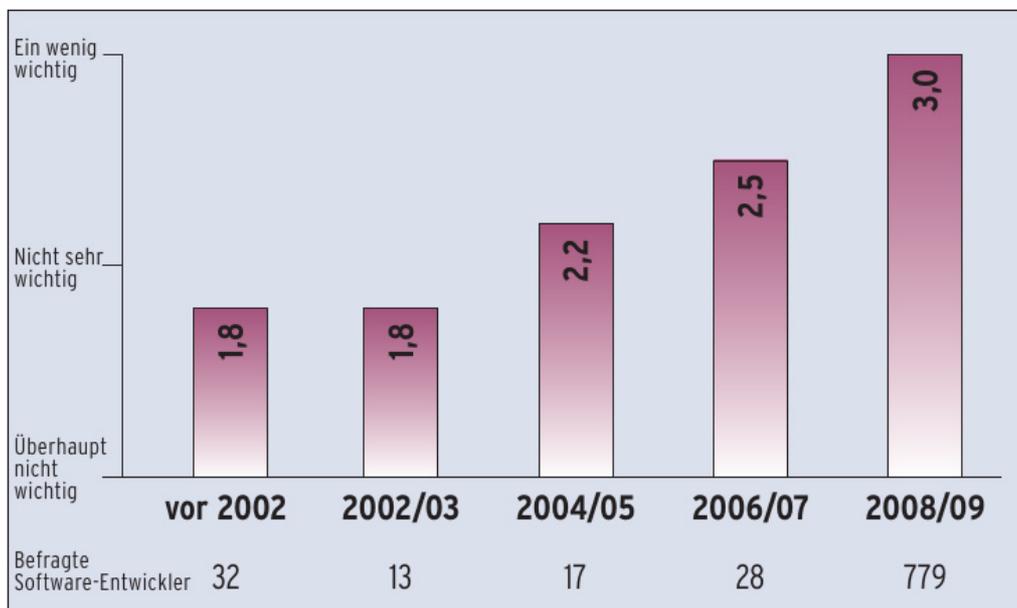


Abbildung 2: Die Einstellung von Entwicklern zum Open-Source-Reusing hat sich über die Zeit gewandelt.

Neben dem Zeitverlauf scheinen Erfahrungen und berufliche Stellung die Einstellung zum Thema zu beeinflussen. So lässt sich feststellen, dass professionelle Softwareentwickler, die über eigene Erfahrungen in Open-Source-Projekten verfügen, solchen Code auch professionell bedeutsamer einschätzen. Möglicherweise tun sich solche Entwickler leichter damit, passenden Code für ihre Projekte zu finden und zu evaluieren.

Interessanterweise beeinflusst auch die eigene Tätigkeit im Unternehmen die Haltung zu Open Source. Softwarearchitekten und Projektmanager zeigen sich gegenüber der Wiederverwendung von Code aus dem Internet aufgeschlossener als Programmierer. Deren Einschätzung wiederum liegt höher als die von Testern, Datenbankentwicklern und Systemanalysten. Dass sich die drei zuletzt genannten Rollen wenig mit dem konkreten Implementieren beschäftigen, macht die geringere Bedeutung plausibel.

Der Unterschied zwischen Softwarearchitekten und Projektmanagern auf der einen Seite und Programmierern auf der anderen mag daran liegen, dass die erste Gruppe über die Beeinflussung der Architektur der zu entwickelnden Software über höhere Freiheitsgrade

beim Einbinden von existierendem Code verfügt als Programmierer, für die die grundlegende Architektur der Software, an der sie arbeiten, bereits von außen vorgegeben ist.

Als Drittes zeitigt die primär verwendete Programmiersprache einen Einfluss. Befragte, die hauptsächlich mit Ruby oder Python arbeiten, erachten wiederverwendeten Code als besonders bedeutsam. Dies mag daran liegen, dass es beide Sprachen besonders einfach machen, fremden Code einzubinden, sogar solchen aus anderen Sprachen. Perl, Javascript, Java und PHP bilden eine zweite Gruppe, erst dann folgen mit C und C++ die traditionellen Sprachen. Entwickler, die mit Fortran, Visual Basic, C# oder Pascal arbeiten, legen den wenigsten Wert auf Code aus dem Internet. Kulturelle Unterschiede scheint es hingegen nicht zu geben, die Entwickler aus Europa, Nord- und Südamerika oder Asien messen quelloffenem Code etwa die gleiche Bedeutung zu.

Kaum offiziell unterstützt

Angesichts der insgesamt hohen Bedeutung der Wiederverwendung von Open-Source-Code in kommerziellen Entwicklungsprojekten ist es erstaunlich, dass der Großteil der befragten 2009 aktiven Softwareentwickler zu diesem Thema nie offiziell geschult worden ist. Auf die Frage, aus welchen Quellen sie sich zu den Vorteilen und Risiken bei dem Thema informieren, nennen zwei Drittel das Internet und knapp die Hälfte Freunde und Kollegen (siehe [Abbildung 3](#)). Ein Viertel der Befragten gab an, überhaupt keine Informationen erhalten zu haben. Aus offiziellen Quellen wie dem eigenen Unternehmen oder einer Ausbildungseinrichtung hatten nur 21 beziehungsweise 16 Prozent Informationen erhalten.

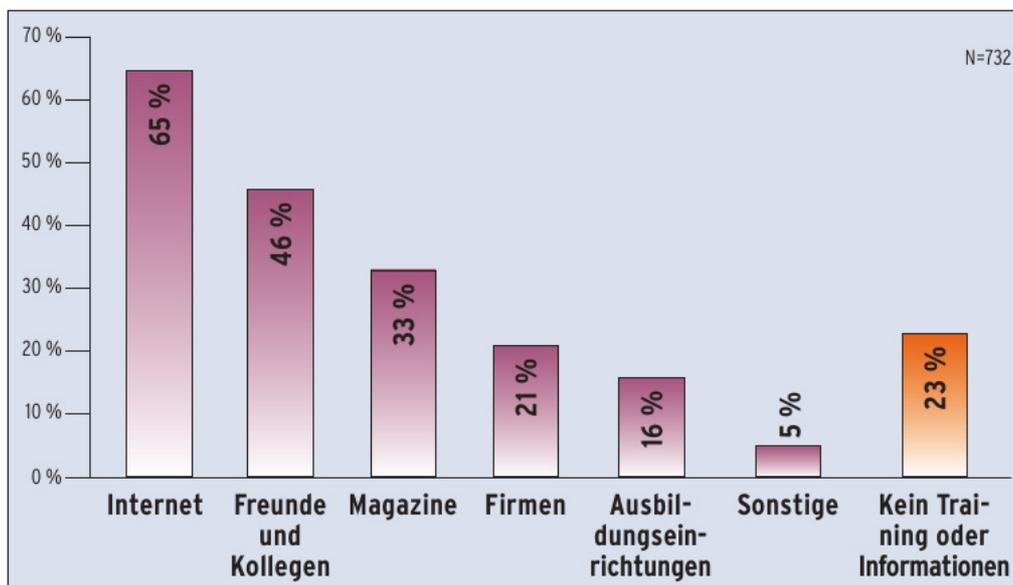


Abbildung 3: Benutzte Informationsquellen zu verwendbarem Code aus dem Internet.

Rechtliche Risiken

Dass jedermann Code aus dem Internet unentgeltlich laden kann, besagt nicht, dass keine Bedingungen daran geknüpft sind. Geschriebener Code unterliegt automatisch dem Urheberrecht, und dieses geht von einer exklusiven Nutzung durch den Urheber aus. Der kann Dritten jedoch die Nutzung und die Vervielfältigung des Codes durch Lizenzen gewähren [8], indem er Rechte abtritt [9].

Dieses Umstandes bedienen sich auch die Open-Source-Lizenzen, die über ihre Lizenztexte definieren, unter welchen Bedingungen der so beauftragte Code verwendbar ist. Die Open Source Initiative [10] zählt alleine 69 Lizenzen, die der Definition Open Source folgen. Daneben gibt es eine Menge von Lizenzen, die zwar nicht den Kriterien der Open-Source-Definition genügen, jedoch eine kostenlose Nutzung des Codes erlauben.

Die unterschiedlichen Lizenzen erteilen spezielle Rechte und stellen unterschiedliche Bedingungen [9]. Eine gängige Bedingung ist beispielsweise das Erwähnen des Autors im Code und in der Dokumentation zur Software oder das Hinzufügen des originalen Lizenztexts. Problematisch für manche Unternehmen scheint insbesondere die Reziprozität zu sein, die manche Lizenzen, unter ihnen die GPLv2, als Bedingung enthalten. Diese sieht vor, dass Software, die eng mit dem freien Code interagiert, ebenfalls dieser Lizenz unterstellt gehört [9].

Bei der klassischen Wiederverwendung von Code lassen sich rechtliche Themen in der Regel einfach handhaben: Sofern es sich um innerhalb der eigenen Firma entwickelten Code handelt, liegen die Rechte schon beim Unternehmen. Wird der Code käuflich erworben, regelt meist der Kaufvertrag die Lizenzfragen. Bei Code aus dem Internet scheint die Gefahr größer zu sein, dass, vor allem bei knappen Terminen, die rechtlichen Aspekte nicht die nötige Beachtung finden.

Knackpunkt Reziprozität

Für Unternehmen stellen Verletzungen der Lizenzen von wiederverwendetem Code ein signifikantes Risiko dar, insbesondere, wenn es sich um Reziprozität fordernde Lizenzen handelt. VMware beispielweise weist im Risiko-Teil eines Quartalsberichts [11] explizit darauf hin: "... use of 'open source' software in our products could negatively affect our ability to sell our products and subject us to possible litigation".

Da Entwickler Code aus dem Internet unbürokratisch herunterladen können, spielen die betreffenden rechtlichen Rahmenbedingungen für sie eine bedeutendere Rolle als zu Zeiten der klassischen Zweitnutzung von firmenintern erstelltem oder zugekauftem Code. Das macht den juristischen Wissensstand professioneller Softwareentwickler interessant.

Gemäß ihrer Selbsteinschätzung sind sich die 2009 aktiven befragten Softwareentwickler der rechtlichen Rahmenbedingungen, die mit der Wiederverwendung von Code aus dem Internet einhergehen, bewusst. 50 Prozent sehen sich selbst als mit dem Thema "vertraut" und behaupteten, dass sie nahezu alle möglichen Fallstricke kennen und mit ihnen umzugehen wissen. 16 Prozent gaben sogar an, Experten auf diesem Feld zu sein, die von anderen dazu gefragt werden. Die Gegenüberstellung dieser Selbsteinschätzung mit den Ergebnissen eines Quiz mit fünf Fragen zu möglichen rechtlichen Problemen bei der Wiederverwendung von Code aus dem Internet zeichnet jedoch ein anderes Bild (siehe [Kasten "Lizenz-Quiz"](#)).

Lizenz-Quiz

Um die Kenntnisse der Umfrageteilnehmer zum Thema Softwarelizenzen nicht allein ihrer Selbsteinschätzung zu überlassen, haben die Forscher mit den Probanden ein Multiple-Choice-Quiz veranstaltet. Hier nun die fünf Fragen, die Prozentzahlen

spiegeln die Antworten der Befragten wider. Per Kreuzchen im Kästchen kann jeder Linux-Magazin-Leser selbst sein Wissen testen.

.1. Welche Open-Source-Lizenz verlangt, dass so regulierter Code auf private oder akademische Software-Projekten beschränkt bleibt?

- GPL (5 %)
- BSD License (10 %)
- Keine der genannten (68 %)
- Weiß nicht (17 %)

2. Angenommen jemand schreibt ein proprietäres Programm und bettet darin Open-Source-Code ein. Welche Lizenz erfordert es unter Umständen, dass der Entwickler das ganze Programm unter die betreffende Open-Source-Lizenz stellen muss?

- GPL (56 %)
- BSD License (1 %)
- Mozilla Public License (MPL) (1 %)
- Sowohl GPL als auch MPL (27 %)
- Keine der genannten (1 %)
- Weiß nicht (14 %)

3. Falls Open-Source-Code ein Patent verletzt, gegen wen kann der Patenteigentümer Ansprüche geltend machen?

- Nur den gegen den Entwickler des Codes (6 %)
- Gegen den Entwickler und Dritte, die den Code integrieren (52 %)
- Niemanden, weil Open-Source-Lizenzen Patente ausschließen (3 %)
- Weiß nicht (39 %)

4. Welche der Open-Source-Lizenzen verlangt, dass solchen Code enthaltenden Produkte den Lizenztext mit ausliefern?

- GPL (16 %)
- BSD License (3 %)
- MPL (1 %)
- GPL, BSD und MPL (60 %)
- Keine der genannten (1 %)
- Weiß nicht (19 %)

5. Jemand veröffentlicht ein Code-Snippet in einer Newsgroup oder auf einer Tutorial-Webseite. Unter welchen Umständen kann ein Dritter das Snippet rechtlich gefahrlos integrieren?

- Wenn der Veröffentlichler keine Verpflichtungen nennt, die mit der Verwendung des Snippets einhergehen (10 %)
- Wenn der Veröffentlichler explizit darauf hinweist, dass er keine Verpflichtungen an die Verwendung des Snippets knüpft (39 %)
- Wenn das Snippet nicht Teil eines kompletten Programms ist (1 %)
- Wenn alle drei genannten Bedingungen erfüllt sind (16 %)
- Keine der genannten Bedingungen reicht aus (19 %)
- Weiß nicht (15 %)

Wer die Auflösung des Rätsels wissen will, schaut ans Ende des Artikels. [Text gegenüber der Printausgabe geändert]

Selbsteinschätzung und Wirklichkeit

Mehr als ein Drittel der 2009 aktiven Entwickler kam im Quiz auf weniger als die Hälfte der Punkte, und nur 3 Prozent beantworteten alle fünf Fragen korrekt. Im Durchschnitt kamen die Softwareentwickler, die sich selbst als "Sehr vertraut" mit den rechtlichen Themen bezeichnen, nur auf 3,1 von 5 Punkten (siehe [Abbildung 4](#)). Für Unternehmen erscheint es darum geboten, angestellte Softwareentwickler anzuhalten, sich ihres eigenen juristischen Kenntnisstands bewusst zu werden und sich mit rechtlichen Themen an tatsächliche Experten zu wenden. Wie zu erwarten, haben Softwareentwickler mit Erfahrungen in öffentlichen Open-Source-Projekten einen höheren tatsächlichen Wissenstand als Kollegen ohne solche Erfahrungen.

Es erscheint für Unternehmen zudem sinnvoll, klare Regeln zu definieren, wie sich ihre Entwickler bei der Wiederverwendung von Code zu verhalten haben. Von den 2009 aktiven Softwareentwicklern arbeiteten jedoch nur knapp über ein Drittel für ein Unternehmen mit solchen festen Regeln. Wie zu erwarten verfügten größere Unternehmen und Unternehmen, deren Hauptgeschäftstätigkeit die Entwicklung von Software ist, mit einer höheren Wahrscheinlichkeit über derartige Regeln als Firmen, die Software beispielsweise nur als Teil ihrer eigentlichen Hauptprodukte entwickeln.

Direkt darauf angesprochen, ob sie in der Vergangenheit möglicherweise Code aus dem Internet in einer Art und Weise wiederverwendet haben, die rechtliche Risiken mit sich brachte, gaben circa 15 Prozent der Befragten an, sich bereits mindestens einmal nicht ausführlich über Bedingungen des recycelten Codes informiert oder gar bekannte Ansprüche wissentlich ignoriert zu haben.

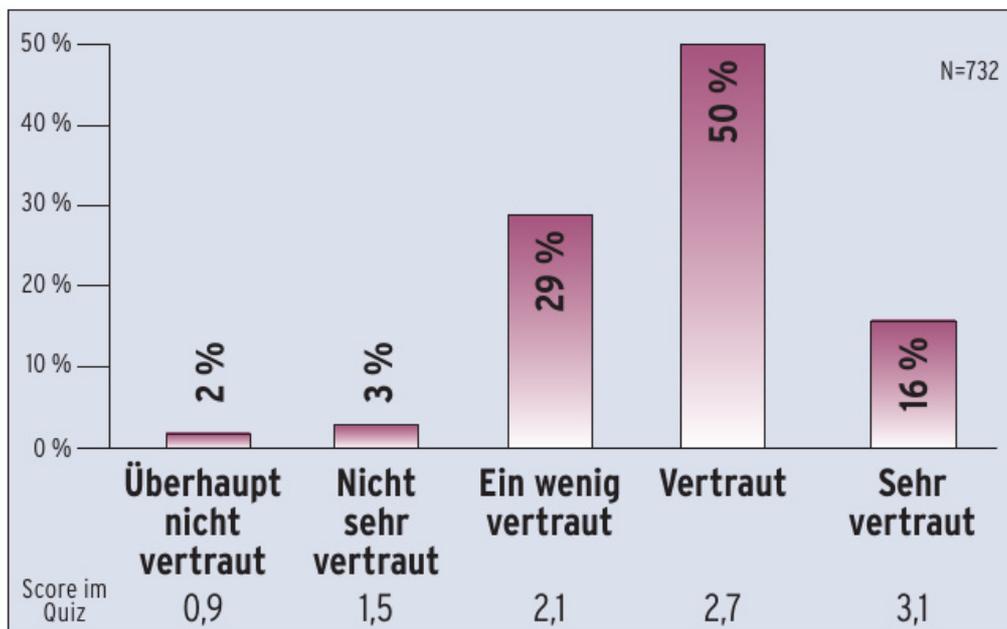


Abbildung 4: Selbsteinschätzung der Entwicklern zu Erfahrungen bezüglich rechtlicher Nachnutzungsrisiken. Die Zeile unten gibt dagegen die erzielten Ergebnisse im Lizenz-Quiz wieder.

Zusammenfassung

Code aus dem Internet in der kommerziellen Entwicklung zu benutzen, ist ein Phänomen, das spätestens 2003 Fahrt aufgenommen hat und heute in vielen Projekten eine übliche Praxis ist. Die in diesem Artikel beleuchtete Umfrage hat ergeben, dass mehr als 50

Prozent der beruflichen Entwickler Code aus dem Internet für ihre Arbeit als "ein wenig wichtig" für ihre Arbeit ansahen. Nur 12 Prozent reagieren komplett immun gegen derartigen Code.

Interessanterweise hatte nur ein kleiner Teil der Befragten während der Ausbildung oder von seinem Arbeitgeber Nutzungsinformationen zu Wiederverwendung fremden Codes erhalten. Der Großteil bezog sein Wissen aus dem Internet oder von Freunden und Kollegen. Dennoch schätzten die meisten Entwickler ihr Wissen zu Lizenzen als hoch ein – viel zu hoch, wie die Auswertung des hier vorgestellten Quiz nahe legt. Auch in den Firmen scheinen die rechtlichen Risiken, die sich beim Praktizieren des Open-Source-Basars ergeben, nicht angekommen zu sein. 2009 verfügten nur knapp ein Drittel der Unternehmen über entsprechende Regeln.

Viele Unternehmen, die Software entwickeln, müssen davon ausgehen, dass ihr Code Fragmente von Code aus dem Internet enthält. Quelloffenen Code nachzunutzen hilft dabei, Software schneller und wirtschaftlicher entwickeln. Auf lange Sicht profitieren Unternehmen jedoch nur, wenn sie dabei Lizenzverletzungen unterbinden – durch Sensibilisieren und Schulen der Entwickler und indem sie eindeutige Regeln und klare Prozesse definieren. *(jk)*

Auflösung Lizenz-Quiz

1. Keine der genannten
2. Sowohl GPL als auch MPL
3. Gegen den Entwickler und Dritte, die den Code integriert haben
4. GPL, BSD und MPL
5. Keine der genannten Bedingungen reicht

Infos

1. Black Duck Software: [<http://www.blackducksoftware.com>]
2. Y. E. Kim, E. A. Stohr, "Software Reuse – Survey and Research Directions" (1998): Journal of Management Information Systems 14(4), S. 113-147
3. A. Lynex, P. J. Layzell, "Organisational Considerations for Software Reuse" (1998): Annals of Software Engineering 5(1), S. 105-124
4. M. Sojer, "Code wiederverwenden in öffentlichen Open-Source-Projekten": Linux-Magazin 04/12, S. 84
5. G. Schryen, "Is Open Source Security a Myth?" (2011): Communications of the ACM 54(5), S. 130-140
6. I. Stamelos, L. Angelis, A. Oikonomou, G. L. Bleris, "Code Quality Analysis in Open Source Software Development" (2002): Information Systems Journal 12(1), S. 43-60

7. M. Sojer, "Reusing Open Source Code" (2010): Gabler Verlag, Wiesbaden
 8. L. Rosen, "Open Source Licensing: Software Freedom and Intellectual Property Law" (2004): Prentice-Hall, Englewood Cliffs, NJ.
 9. B. Fitzgerald, G. Bassett, "Legal Issues Relating to Free and Open Source Software", In: B. Fitzgerald, G. Bassett (Ed.), "Legal Issues Relating to Free and Open Source Software" (2005), Queensland University of Technology, Brisbane, QLD, S. 11-36
 10. Open Source Initiative: [<http://opensource.org>]
 11. VMware, Inc., "Form 10-Q, Filed 08/02/12 for the Period Ending 06/30/12": [<http://ir.vmware.com/common/download/sec.cfm?companyid=VMW&fid=1124610-12-21&CIK=1124610>]
-

Der Autor

Dr. Manuel Sojer hat 2010 am Dr. Theo Schöller-Stiftungslehrstuhl für Technologie- und Innovationsmanagement der Technischen Universität München bei Prof. Dr. Joachim Henkel zur Wiederverwendung von Open-Source-Code promoviert.

© 2012 COMPUTEC MEDIA GmbH

Schwesterpublikationen:

[[Linux-Magazin](#)] [[LinuxUser](#)] [[Raspberry Pi Geek](#)] [[Linux-Community](#)] [[Computec Academy](#)] [[Golem.de](#)]