2019 State of the Software Supply Chain

The 5th annual report on global open source software development

presented by

in partnership with
# Table of Contents

Introduction ................................................................................. 3
Infographic .................................................................................. 4

## CHAPTER 1: Global Supply of Open Source .......................... 5
1.1 Supply of Open Source is Massive ........................................ 6
1.2 Supply of Open Source is Expanding Rapidly .................. 7
1.3 Suppliers, Components and Releases ................................. 7

## CHAPTER 2: Global Demand for Open Source .................... 8
2.1 Accelerating Demand for Open Source Libraries ............... 9
2.2 Automated Pipelines and DevOps Are Key Drivers .......... 10

## CHAPTER 3: Exemplary Project Teams ............................... 11
3.1 Research Goals ...................................................................... 12
3.2 Time to Remediate Vulnerabilities ..................................... 13
3.3 Time to Update Dependencies ............................................ 15
3.4 Stale Dependencies ............................................................ 16
3.5 Exploring the Link Between MTTR and MTTU ................. 16
3.6 Hypothesis Testing ............................................................. 18
3.7 Finding Different Behavioral Groups ................................. 22
3.8 Guidance for Open Source Project Owners and Contributors ........................................................................ 24
3.9 Guidance for Enterprise Development Teams .................. 24

## CHAPTER 4: Exemplary Dev Teams .................................... 26
4.1 The Enterprise Continues to Accelerate ............................ 27
4.2 Analysis of 12,000 Large Enterprises ............................... 27
4.3 Component Releases Make Up 85% of a Modern Application ........................................................................ 28
4.4 Characteristics of Exemplary Development Teams ............ 29
4.5 Rewards for Exemplary Development Teams .................. 34

## CHAPTER 5: The Changing Landscape .............................. 35
5.1 Deming Emphasizes Building Quality In .......................... 36
5.2 Tracing Vulnerable Component Release Downloads Across Software Supply Chains ........................................... 37
5.3 Adversaries Increasingly Target Open Source Components ........................................................................ 38
5.4 Government and Industry Apply New Standards to Secure Software Development ........................................ 42

Conclusion .................................................................................. 44
Sources ....................................................................................... 45

Appendix A ............................................................................... 46
Acknowledgments ....................................................................... 46
About the Analysis ...................................................................... 46

Appendix B ............................................................................... 47
Appendix C ............................................................................... 49
Appendix D ............................................................................... 50
Introduction

Now in its fifth year, Sonatype’s annual State of the Software Supply Chain Report examines the rapidly expanding supply and continued exponential growth in consumption of open source components. Our research also reveals best practices exhibited by exemplary open source software projects and exemplary commercial application development teams.

This year, for the first time, we’ve collaborated with research partners Gene Kim from IT Revolution and Dr. Stephen Magill, Principal Scientist at Galois and CEO of MuseDev, to objectively examine and empirically document, release patterns and hygiene practices across 36,000 open source project teams and 3.7 million open source releases.

Separately, we observed 12,000 commercial engineering teams to document their consumption of open source and third party libraries. We also conducted two surveys, with a combined participation of over 6,200 development professionals, to better understand the current state of DevSecOps and software supply chain management practices. We compared teams with, and without, automated open source governance capabilities (in an attempt) to reveal the baseline benefit of building applications that utilize higher quality open source components.

The 2019 State of the Software Supply Chain Report blends a broad set of public and proprietary data with survey results, expert research, and analysis to reveal the following:

- 75% growth in supply of open source component releases over the past two years (Chapter 1)
- 68% year over year growth in download requests from the Central Repository to 146 billion (Chapter 2)
- 18x faster median time to update dependencies for exemplary open source components (Chapter 3)
- 55% reduction in the use of vulnerable open source component releases within managed software supply chains (Chapter 4)
- 71% increase in confirmed or suspected open source related breaches since 2014 (Chapter 5)

At Sonatype, we’ve long been active contributors to, and maintainers of, numerous open source efforts, including Apache Maven and Nexus Repository Manager. Since 2005, we’ve curated and operated the Central Repository, which last year alone serviced 146 billion download requests for component releases from developers around the world. Commercially, our interest lies in helping developers and organizations accelerate innovation and minimize risk by continuously sourcing third-party libraries from the highest quality open source projects.

Together with our partners, we are proud to share this research. We hope that you find it valuable.
51% of JavaScript components have a known security vulnerability

The top 5% of projects remediate security vulnerabilities within 21 days

12x more likely
313,000 average enterprise downloads of OSS components per year

3.4x faster at remediating known vulnerabilities

10x more likely to schedule dependency updates

18x faster at updating dependencies

146B Java download requests in 2018, represented a 68% year over year growth

1 in 10 downloads of Java component releases have known vulnerabilities

1 in 10

71% increase in open source related breaches over the past five years

Secure Dev Practices are 9.3x more likely to proactively remove troublesome dependencies

Exemplary Projects are 3.4x faster at remediating known vulnerabilities

Exemplary Dev Teams experience a 55% reduction in the use of vulnerable OSS component releases in automated environments

Exemplary Dev Teams are 12x more likely to have automated tools to manage open source dependencies

Exemplary Dev Teams are 10x more likely to schedule dependency updates

Exemplary Projects are 18x faster at updating dependencies

313,000
average enterprise downloads of OSS components per year

PCI introduces new standards for development teams using open source components

21,448 new open source releases per day

Exemplary Dev Teams are 18x faster at updating dependencies

Exemplary Projects are 71% increase in open source related breaches over the past five years
CHAPTER 1

Global Supply of Open Source

A World of Infinite Choice
1.1 Supply of Open Source is Massive

There are now more than 3.7 million unique Java open source software component releases in the Central Repository, 800,000 unique JavaScript packages in npm, 1.2 million unique Python component releases housed in the PyPI repository, and 1.6 million .NET component releases in the NuGet Gallery. There are also more than 2.2 million containerized applications housed in Docker Hub — up from 900,000 the previous year.

This massive supply of software parts is rapidly and organically expanding due to constant innovations and regular versioning of existing components. These new versions not only offer enhanced features, but also provide improved performance, bug fixes, and security patches.

1.2 Supply of Open Source is Expanding Rapidly

Sonatype’s study across several open source component ecosystems reveals number of releases housed within public repositories increased from 16.6 million to 28.4 million from January 2018 through today. On average, developers had access to more than 21,448 new open source component releases every day, since the beginning of 2018.

**KEY POINT**

- On average, developers had access to more than 21,448 new open source component releases every day, since the beginning of 2018.
While Java components continue to dominate by sheer number of component releases available, package types like npm and crates.io demonstrated the highest growth rates. npm packages experienced 109% growth and now total 836,000 component releases. Newcomer crates.io (Rust) increased 213% during the period and now offers more than 25,000 component releases (see FIGURE 1A).

Open source growth is robust across numerous ecosystems, but npm has grown particularly fast due to JavaScript’s emergence as a universal web application programming language. For JavaScript developers looking for a library, tool, or adapter, odds are very good that someone else has already created it and published it to npm where it can easily be borrowed. According to the stewards of the npm repository, “Every week roughly 160 people publish their first package in the registry.”

New component versions are released for a variety of reasons, including supporting new functionality, fixing defects, or supporting a new API. New releases may also address non-feature-related concerns, such as improving performance, adding or updating their dependencies, or remediating security vulnerabilities.

Although brand new projects are constantly being created and introduced, growth in open source supply is driven mostly by new versions of existing projects being published. While the growth in supply fuels rapid innovation, it does pose significant questions for organizations wanting to better manage their software supply chains:

- How often do projects publish new versions?
- Do certain projects release updates more frequently?
- Do other projects release updates less frequently?
- What are the implications?
- Who are the best component suppliers?

This year, State of the Software Supply Chain researchers set out to answer these questions and many more.

1.3 Suppliers, Components and Releases

To match terminology in the previous State of the Software Supply Chain reports and provide consistency across research findings, we will follow the Maven terminology using the following definitions:

- A **supplier** is a Maven Group (e.g., org.apache.httpcomponents)
- A **component** is a Maven Group and Artifact (e.g., org.apache.httpcomponents, httpclient)
- A **component release** is a specific Maven Group, Artifact and Version (e.g., org.apache.httpcomponents, httpclient, 4.5.6). For other ecosystems, we sometimes refer to releases as packages.

Fueling Rapid Innovation Consumption of open source is so vast that most organizations cannot identify how many components are entering into their software supply chains, how those components are flowing through development lifecycles, the relative quality and security of those components, or which components exist within production applications.

According to International Data Corporation (IDC), “there were 22.30 million software developers in the world at the outset of 2018. IDC estimated that 11.65 million are full-time developers, 6.35 million are part-time developers, and 4.30 million are nonprofessional developers.” In 2018, developers around the world consumed hundreds of billions of open source software component releases.
CHAPTER 2

Global Demand for Open Source

Fueling Rapid Innovation
2.1 Accelerating Demand for Open Source Libraries

The growing demand for innovation has accelerated implementations of automated software development pipelines while also driving open source consumption to new heights across all major ecosystems.

2.1.1 Demand for Java
In 2018, the aggregate number of download requests for Java component releases from the Central Repository grew 68% year over year to 146 billion. With an estimated 12 million Java developers around the world, this equates to 12,166 per person.\(^7\)

2.1.2 Demand for JavaScript
While growth in demand for Java components is remarkable, a view into JavaScript package downloads demonstrates even greater growth in developer demand. In 2018, average weekly npm package downloads increased from approximately 3.5 billion to 10 billion — an increase of 185%.\(^8\) To add further context, there are an estimated 9.7 million JavaScript developers in the world — meaning the average JavaScript developer is sourcing 1,030 packages per week or 53,608 packages per year.\(^9\)

2.1.3 Others
Other public repositories have experienced similar levels of developer demand. For example, downloads of RubyGems packages have grown from 8 billion to 33 billion over the past three years.\(^10\) NuGet package downloads have increased from 756 million in 2015 to an annual run rate of 16.2 billion in 2019.\(^11\)
2.2 Automated Pipelines and DevOps Are Key Drivers

Exponential growth in the consumption of open source component releases and containers is a proxy for the adoption of automated software development tools and DevOps pipelines. Automated tooling can generate hundreds or thousands of download requests per build.

In the context of software supply chain management, each download equates to a procurement effort by development teams. Each open source software component release is chosen from an OSS project that acts as a supplier to developers who assemble tens, hundreds, and sometimes thousands of component releases into a finished application.

A view into JavaScript component downloads demonstrates even greater growth in developer demand. In 2018, average weekly npm package downloads increased from approximately 3.5 billion to 10 billion — an increase of 185%.

FIG. 2B JavaScript Package Downloads, Rolling Weekly Average 2013 – 2019
SOURCE: NPM INC., LAURIE VOSS (@SELD0)
CHAPTER 3

Exemplary Project Teams

Open Source Projects Are Not Created Equal
3.1 Research Goals

We wanted to better understand the health and habits of the open source component ecosystem, and how software developers choose which OSS components to use in their own projects. To do this, we studied all the Java artifacts stored in The Central Repository (often referred to as “Maven Central”). At the time of this writing, The Central Repository has over 266,000 unique components with over 3.7 million component releases.

The research set out to answer:

- Do differences exist in how effectively OSS projects update their dependencies and fix vulnerabilities? Are there exemplary components that do this better than others?
- Are exemplary components more widely-used than “non-exemplary” components?
- What factors correlate with exemplary components?
- What advice can be offered to producers of OSS components and the developers that consume them?

For the purpose of this study, we restricted our analysis to components that met six qualifying criteria (see FIGURE 3A), resulting in a data set of 36,203 components. Given this data set, we examined a number of attributes to measure responsiveness to security vulnerabilities and determine what properties exemplary component project teams shared. In addition to the security relevant attributes — described in the next section — the primary attributes tracked were:

Number of Dependencies: the maximum count of dependencies for any given component across all versions in the study period, as measured by the dependencies in the Maven pom.xml file.

Stale Dependencies (fewer is better): the average percentage of out-of-date component dependencies (i.e., a newer version has been released) present when the component has a new release.

Release Period (shorter is better): average time in days each component version spends as the “current” release. A shorter average release period equates to more frequent releases.

Popularity: the average number of downloads per day from The Central Repository.

SCM Commits per Month: average number of commits to the GitHub repo per month — a measure of development activity (i.e., developer velocity).

FIG. 3A Constructing the Study Dataset (N = 36,203)

- N = 266,170 Components were published in The Central Repository.
- N = 168,231 Components had at least two version releases in the last five years.
- N = 101,252 Components were part of the “open source software supply chain” (e.g., they are or they have a dependency).
- N = 100,643 Components follow the Maven standard for versioning guidance. (e.g., correct use of numeric version strings, components separated by dots.)
- N = 76,795 Components have dependencies satisfying all of the above.
- N = 36,203 Components have updated a dependency at least once.
**Developer Team Size:** as measured by the average number of unique developers committing each month.

**Presence of Continuous Integration (CI):** as measured by the detection of any CI-related configuration files in the source code repository (e.g., Travis, Jenkins, CircleCI, etc.).

**Support Type:** support for the component comes from an open source foundation, a commercial organization, or is not officially supported by any organization (e.g., a personal project).

To gather these attributes, we augmented the data as follows:

- **CHAOS:** using the perceval utility to gather GitHub commit data, we gathered the number of commits per month for twelve months, as well as the number of unique developers committing during each month.
- **Libraries.io:** using this dataset, we were able to gather the number of GitHub stars, forks, and pull requests.
- **Sonatype Nexus IQ Server:** using Sonatype data, we are able to gather vulnerability information and a derived component popularity score (based on how frequently components are seen by the Nexus IQ repository scanning service).

### 3.2 Time to Remediate Vulnerabilities

The first outcome metric we focused on was time to remediate (TTR). TTR is the time required for a component team to remediate any security vulnerabilities reported against their dependencies. For a vulnerable dependency, remediation occurs when it is upgraded to a new version that fixes the vulnerability. We combined data on component updates with data on vulnerabilities, and tagged updates as *security relevant* if there was a known vulnerability targeting the old version at the time the new version was released. We then measure mean time to remediate (MTTR) as the average time required for a component to adopt security relevant updates to dependencies.

Note that the TTR clock starts when a component version fixing a vulnerability is released. Specifically, the TTR clock is not started when the vulnerability is reported or published; early efforts used vulnerability publish times from various sources as the TTR starting time led to many problems. For example, responsible disclosure often delays the publishing of the vulnerability until developers have released a fixed component version, or vulnerabilities may have CVE numbers associated with the date they were reserved rather than published. These factors make it difficult to obtain precise dates of when component teams were advised of vulnerabilities. On the other hand, data on which components fixed a known vulnerability with a new version release is widespread and much more reliable.

For developers wanting to use more secure components, those exhibiting faster MTTR are desirable. Figure 3B (we cropped the x-axis at the 95th percentile) shows a histogram of MTTR across all components, demonstrating the long tail of components that apply security updates very slowly, often years after they are released. We observed:

- In the study, 47% of components — after releasing a new version — had a vulnerability discovered in one of its dependencies, during the period in which that version was current. (N = 36,203)
- The median TTR was 180 days (similar to numbers reported in previous versions of the State of the Software Supply Chain Report). The top 5% of components remediated vulnerabilities within 21 days.

**KEY POINTS**

- TTR is the time required for a component team to remediate any security vulnerabilities reported against their dependencies.
- For developers wanting to use more secure components, those exhibiting faster MTTR are desirable.
- In the study, 47% of components — after releasing a new version — had a vulnerability discovered in one of its dependencies, during the period in which that version was current.
The TTR has an extremely “long tail” — the 95% percentile occurs at 1,302 days (3.5 years), with the maximum TTR at 3,388 days (9.3 years). The mean of the TTR is 326 days.

59% of components have had at least one non-up-to-date and known vulnerable dependency at the time of release (e.g., a newer dependency was available that fixed a known vulnerability).

During the period studied, many components had to remediate a security vulnerability caused by a dependency, but over half of the sample set did not have vulnerabilities in any of their direct dependencies. Furthermore, we focused further study on components that were published on GitHub, where we could gather developer activity metrics. When we restrict the population to GitHub-hosted components, the portion that had to deal with a security-relevant dependency update dropped to 16%. In other words, our TTR data set was much smaller than our component data set. This, combined with the hypothesized correlation between median time to update (MTTU) and MTTR, motivated us to look at MTTU as a primary project health metric.

**KEY POINTS**

- 59% of components have released at least one version that contained a dependency with a known vulnerability.
- Over half of the sample set did not have vulnerabilities in any of their direct dependencies.
3.3 Time to Update Dependencies

To enable a broader analysis of dependency update hygiene, we defined a time to update (TTU) attribute for all components in our dataset.

For each component in our data set, we constructed the dependency graph of the set of components (and their versions) that each component release depends on.

Then for the given period, for every component, whenever a new version of one of its dependencies was released, we measured the time required for the component to update to the newer dependency. We then computed the median of all those time to update data points to obtain the MTTU metric.

FIGURE 3C shows three of our key metrics — time to update (TTU), time to remediate (TTR), and stale dependencies. Suppose:

- Component C depends on Component A and B.
- Component B (version 2.2) has a vulnerability published against it that is fixed in version 2.3.
- The release of B (version 2.3) starts the Component C TTR and TTU clock.
- When Component C updates Component B from version 2.3 to 2.4, the clock stops and we record the TTR and TTU.
- Since the release of A (version 2.4) is not security relevant (no vulnerability known against A), upgrading this component only contributes to TTU.
- When C (version 2.2) is released it is using an old version of A (2.2 rather than 2.3); this causes A to be regarded as a stale dependency for the release of C.

There are several reasons we suspected that MTTU would be an effective indicator of MTTR performance. One of the most important was phrased by Jeremy Long, founder of the OWASP Dependency Check project who recommends the best security patching strategy is to remain current on all dependencies. Long speculates that “only 25% of organizations report vulnerabilities to users, and only 10% of vulnerabilities are reported as Common Vulnerabilities and Exposures (CVE).” Furthermore, the publication of a CVE is often for a vulnerability that was fixed in an earlier version of a component.

As an example, Long cites a security vulnerability discovered in PrimeFaces — a Java UI framework. The PrimeFaces project became aware of the vulnerability and fixed it in February 2016. A CVE for this vulnerability (CVE-2017-1000486) was subsequently assigned in 2017. Then, the CVE was published into the national catalogue on

KEY POINT

- Jeremy Long, founder of the OWASP Dependency Check project speculates that “only 25% of organizations report vulnerabilities to users, and only 10% of vulnerabilities are reported as Common Vulnerabilities and Exposures (CVE).”
January 3, 2018. Upon the publication of the CVE, crypto-miners actively started exploiting vulnerable versions of the component. Developers who made a practice of updating to the latest released versions of PrimeFaces were less at risk than developers who relied upon publication of a CVE to trigger remediation efforts.

To summarize:

- MTTU data was derivable for all components, whereas MTTR data is more sparse. (N=36,203 for MTTU vs. N = 16,997 for MTTR).
- Fast MTTU makes it more likely that components are already protected when new CVEs are published.
- There is a general sense in the security community that “having better security” is achieved by better technical practices. In this case, better practices means integrating updated dependencies into the daily work of the development team.

We therefore explored TTU in our population to better understand how effectively component teams were updating their dependencies.

### 3.4 Stale Dependencies

Almost every developer has updated their dependencies only to discover that it introduced breaking changes: either compile-time failures, or worse, run-time errors because the dependency functionality has changed. Because of these situations, updating dependencies and patching vulnerabilities becomes an arduous and painful activity, which often leads to developments becoming so behind in updates that upgrading will surely break application functionality (see the survey results in Chapter 4).

To capture the distinction between “fully up to date” and the more conservative “within one version of up to date,” we introduced an attribute called stale dependency percentage. This measures the percentage of dependencies, on average, which are not fully up to date when a component releases a new version. In other words, a component that always releases with all their dependencies up to date will have a perfect stale dependency ratio of 0%.

Projects that stick to the “bleeding edge” of using the latest N or N-1 dependencies will have fast MTTU and a low stale dependency ratio. Projects that tend to stay one version behind will have relatively fast MTTU, but high stale dependency ratio. Projects that do not update frequently will have even slower MTTU and higher stale dependency ratios.

One of the papers we were inspired by was Why and How Java Developers Break APIs that monitored 400 Java libraries for 116 days, and found 282 commits that were breaking changes. Not all of these were in the released components, but it shows that the risk of breaking changes is real, and potentially introduces a huge economic cost of staying current with OSS components.

### 3.5 Exploring the Link Between MTTR and MTTU

Across the entire population, the adoption curve for upgrading dependencies and remediating vulnerabilities are similar, as shown in FIGURE 3D. When comparing MTTR with MTTU for non-security-relevant updates on a per-component basis, we see a correlation between update behavior for security relevant updates (MTTR) and non-security-relevant updates.

#### KEY POINTS

- Fast MTTU makes it more likely that components are already protected when new CVEs are published.
- We introduced an attribute called stale dependency percentage.
- 400 Java libraries were monitored for 116 days. Researchers found 282 commits that were breaking changes.
- When comparing MTTR with MTTU for non-security-relevant updates on a per-component basis, we see a correlation between update behavior for security relevant updates (MTTR) and non-security-relevant updates.
However, we saw a group of components that focused much more on upgrading vulnerable dependencies than applying other updates — achieving good security outcomes while ignoring (deliberately or accidentally) other component updates. The analysis found that 15% of components have a better than average MTTR, while having below average MTTU for non-security relevant updates.

In general, developers staying up to date on dependencies will also stay up to date on security updates, because security updates are a subset of general updates. We observed that many teams follow this practice, exhibiting very similar MTTR and MTTU values. To replicate this practice, security managers can improve vulnerability updating practices by partnering with their development manager counterparts to improve their general dependency management practices.

### 3.6 Hypothesis Testing

Over several months, we came up with a series of hypotheses of what factors might be associated with better upgrade hygiene (i.e., faster MTTU and MTRU) and found that teams that follow the practice of staying up to date on dependencies also stay up to date on security updates. This suggests that improving dependency management practices can lead to better security outcomes.
MTTR). Some of the factors included: number of dependencies, developer commit frequency, number of active developers, continuous integration (CI) usage, and component popularity. We hoped to find specific behaviors associated with exemplary component teams and outcomes. This could provide guidance for other component teams (component producers), and provide developers (consumers in the software supply chain) with a list of attributes to look for in OSS components.

To perform this analysis, we combined our dataset with statistics on development behaviors collected from GitHub via the CHAOSS project’s Perceval tool. Since not all components are hosted on GitHub, this reduced our dataset size to 10,573 components.

**FIG. 3E Correlation Between Security- Relevant and Non-Security-Relevant Update Times**

**KEY POINT**

- Over several months, we came up with a series of hypotheses of what factors might be associated with better upgrade hygiene (i.e., faster MTTU and MTTR).
Our primary hypotheses were:

1. **Popularity**: More popular components will have better update hygiene, due to the pressure to stay secure when a component is widely used.

2. **Number of Dependencies**: Components with fewer dependencies will have better update hygiene, as it is easier to keep current with fewer dependencies.

3. **Size of Team**: Large development teams will have better update hygiene.

4. **Development Activity**: Components that release faster or commit more frequently will have better update hygiene.

5. **Use of Continuous Integration**: Projects that used continuous integration would have better update hygiene.

6. **Institutional Support**: Components that are supported by an open source foundation or commercial entity will have better update hygiene, due to having more resources.

While we found evidence for hypotheses 3 and 4, and weak evidence of 5, the most interesting results were the hypotheses that we could not confirm (1 and 2). For each of our hypotheses regarding component attributes that would be associated with faster MTTU, we obtained the following results.

### 3.6.1 Popularity

**Question**: Do popular projects, as measured by The Central Repository downloads, have better MTTU?

**Findings**: No. When deciding to choose components to use in a development project, popularity is often used as a proxy for quality (i.e., “everyone else is using it, so it must be safe, secure, and reliable”). Across the entire population studied, analysis showed popularity did not correlate with fast MTTU (Pearson correlation coefficient of -0.07, Kendall Tau of -0.1). Even when our researchers selected the most popular components (e.g., the top 10 or 20 percent of components by popularity) and compared their MTTU to the rest of the population, no statistically significant differences in MTTU were found.

### 3.6.2 Number of Dependencies

**Question**: Do fewer dependencies correlate with faster MTTU?

**Findings**: No. We had expected components with fewer dependencies to have better MTTU. One would expect that the difficulty of keeping dependencies up to date should grow as the number of dependencies grow. However, our analysis found very little correlation (Pearson = -0.09). What little correlation there was showed that components with larger dependency counts having slightly faster TTU.

Even more surprising, the analysis showed that components with the most dependencies (i.e., the top 10%) had 39% faster TTU ($p = 1.2e-29$ on Mann-Whitney U test) than the rest of the population. They also had 18% larger development teams than the rest of the population, and this ratio grows as the number of dependencies increase (e.g., at the 95th percentile, the teams were 23% larger). The top 10% by dependency count had 4.8 times more dependencies on average than the bottom 90% and yet had 39% faster TTU. This led us to investigate further the connection between number of dependencies and size of development team. **FIGURE 3F** shows the plot of number of dependencies versus average size of development team (smoothed over a sliding window of 10 data points).

We were surprised and delighted to find that fast TTU is possible even with large numbers of dependencies. It was particularly interesting that the size of the development team tends to increase as the number of dependencies increases, bringing up an interesting question of causation: when development teams grow, does each developer bring in their favorite dependencies, adding to the dependency count? Or is it that as a project grows, new functionality requires new dependencies, which increase the workload, which requires bringing in new developers. These are questions that we hope to address in a future study.

### Key Points

- Across the entire population studied, analysis showed popularity did not correlate with fast MTTU.
- The top 10% by dependency count had 4.8 times more dependencies on average than the bottom 90% and yet had 39% faster TTU.
- Fast TTU is possible even with large numbers of dependencies.
3.6.3 Size of Development Team

**QUESTION:** Do components with more active developers correlate with faster MTTU?

**FINDINGS:** Yes. Selecting directly for size of development team reveals that across all projects, the top 20% of teams by size (11 or more developers contributing per month) have 50% faster MTTU and release 2.6 times more frequently. They are also 37% more likely to be foundation supported. All results are statistically significant (Mann-Whitney U test) and these trends hold for the top 15% and 10% of teams as well (13 and 16 developers involved per month respectively).

3.6.4 Development Activity and Velocity

**QUESTION:** Do components with higher release frequency and higher monthly commits correlate with faster MTTU?

**FINDINGS:** Yes. Clearly, high development activity has the potential to enable a project to stay more up to date. In particular, a project cannot update dependencies without releasing a new version. Therefore, frequent releases are generally required for fast MTTU of dependencies. However, it is also possible to spend development effort solely on new features, bug fixes, etc., and ignore dependency management. Furthermore, development and release velocity have been shown to be highly predictive of project outcomes as noted in the 2018 Accelerate: State of DevOps Report. While we cannot determine causation, we found correlations that match these prior results — the top 20% of teams by commits per month had 26% faster MTTU and 83% faster release frequency. Release frequency itself is strongly correlated with MTTU (Pearson correlation of 0.48). This was the strongest correlation we found among the attributes we considered.

**KEY POINTS**

- The top 20% of teams by size (11 or more developers contributing per month) have 50% faster MTTU and release 2.6 times more frequently.
- The top 20% of teams by commits per month had 26% faster MTTU and 83% faster release frequency. This was the strongest correlation we found among the attributes we considered.
- The top 20% of projects by release frequency have 2.3 times faster TTU on average.

**FIG. 3F More Dependencies Correlate with Larger Development Teams**

(smoothed)
correlation we found among the attributes we considered. The top 20% of projects by release frequency have 2.3 times faster TTU on average.

### 3.6.5 Use of Continuous Integration

**QUESTION:** Does the use of continuous integration correlate with faster MTTU?

**FINDINGS:** No. We were surprised that continuous integration (CI) did not correlate with MTTU. CI was used in the release processes for 68.3% of the components we studied. Usage of CI across all the groups were similar, +/-5%. We speculate that perhaps CI adoption is widespread enough, and that expectations of submitting automated tests with contributed code is high enough, that it is now a mandatory practice for any OSS components. Because we scanned for the presence of CI configuration files in the source code repository, we were able to establish which tools they used. Of the 7,682 repositories, 90.1% were using TravisCI.

### 3.6.6 Institutional Support

**QUESTION:** Is a project supported by an open source foundation or by a commercial entity correlated with faster MTTU?

**FINDINGS:** Yes and no. We hypothesized that the additional resources available to commercially or foundation-supported projects would generally enable better performance. We based commercial support on group IDs that begin with “com,” and foundation support on group IDs associated with multiple components. To improve the labeling, researchers filtered out manually identified outliers such as “com.github,” “org.eclipse,” and “org.webjars,” which all host and aggregate components developed by other groups. The analysis revealed that Foundation-supported projects had 7.4% faster MTTU in general (p = 0.0002), where researchers also observed highly significant differences in team size (32% larger) and commit frequency (77% faster). By comparison, Commercial projects had 8% slower MTTU (p = 0.0001) and 29% slower commit frequency.
3.7 Finding Different Behavioral Groups

Based on what we learned from our hypothesis testing regarding trends in team composition and performance, we examined five sub-populations of the data to characterize differences in approach to open source software development. The identification and characterization of the sub-populations was driven by a combination of automated clustering, domain, and attribute knowledge gained from the hypothesis testing.

3.7.1 Exemplars

We defined Exemplars to be those teams in the fastest 20% by MTTU, and in the best (lowest) 20% by stale dependency count. Exemplars demonstrate statistically significant differences as compared to the rest of the data set in the following attributes:

► 18x faster MTTU
► 6.8x better at releasing components where all dependencies are up to date
► 6x more popular (as measured by average monthly The Central Repository download counts)
► 2x more frequent component releases
► 33% larger development teams
► 4x more likely to be managed by open source foundations than by commercial stewards

Within the Exemplars, there were two groups with significantly different development team sizes:

3.7.1.1 LARGE EXEMPLARS

Large exemplary teams (top 50% by size, with an average of 8.9 developers committing code on at least a monthly basis), commit code frequently, release frequently, and do an excellent job of managing their dependencies. We can see the effect of open source foundation support in this group, as 91% of these projects are associated with an open source foundation.

3.7.1.2 SMALL EXEMPLARS

The smallest 50% of exemplary teams by number of developers have an average of less than two developers, but still manage to run popular, widely used, and high quality projects. However small in team size, they still update dependencies 14 times faster than the rest of the population and stay fastidiously up to date (91% of dependencies are brought up to date with each release).

3.7.2 Laggards

The teams in the bottom 20% in MTTU and stale dependencies are the furthest behind in terms of update hygiene. These teams release infrequently (around twice each year) and take on average almost two years to adopt updates to dependencies. Almost all of their dependencies (98% on average) are out of date, even after a new release. They are generally less popular (downloaded as often as other projects on average). However there are 67 projects in this group that are among the top 10% most downloaded projects from The Central Repository.

3.7.2.1 FEATURES FIRST LAGGARDS

These teams release frequently (top 50%) but otherwise fall into the Laggard category (bottom 20% MTTU and stale dependencies). They have larger than average (29% larger) development teams, but do not prioritize upgrading dependencies. They release a new version every 50 days on average but take an average of 603 days to upgrade dependencies when new versions are released. As a result, 96% of dependencies are out of date at release time. This was a small group, with 2.6% of the population exhibiting this behavior.

3.7.3 Cautious Teams

We checked to see how many teams were in the top 50% with respect to MTTU, but the bottom 20% with respect to stale dependencies. These teams maintain better-than-median update cadence, yet do not immediately adopt new versions of dependencies, choosing instead to wait a few months before moving to a new dependency release. This was not a sizable group, with only 4% of our dataset falling into this category.

continues on page 24
FIG. 3H Cluster Popularity and Release Speed

- Popularity trends up as release speed increases. Exemplars tend to be more popular.
- Features First tend to be fairly popular, despite poor update hygiene.
- A few popular Laggards, but on the whole they are less popular.

Source: 2019 State of the Software Supply Chain Report

- Large Exemplars
- Small Exemplars
- Features First
- Laggards
- Cautious
- None of the Above

Popularity (Maven Central Downloads)

Average Days Between Releases

CHAPTER 3: EXEMPLARY PROJECT TEAMS

2019 STATE OF THE SOFTWARE SUPPLY CHAIN REPORT
3.8 Guidance for Open Source Project Owners and Contributors

Given its association with good security practices and outcomes, we recommend a focus on accelerating and maintaining rapid MTTU. In addition to investing development effort on new features, bug fixes, etc., projects should commit similar resources to dependency management. This means that developers maintaining OSS projects who are considering adding a new dependency, and looking for a metric to guide that choice, would do well to focus on those dependencies with fast MTTU. Since remediating a vulnerable dependency typically involves upgrading to a new dependency version, components with fast TTU values naturally exhibit faster response to dependency vulnerabilities.

To progress comfortably into the status of Exemplar (top 80% of Exemplars), teams should aim for a minimum of four releases annually, and aim to upgrade at least 80% of their dependencies with every release. A higher frequency of dependency updates statistically results in higher quality and more secure code.

3.9 Guidance for Enterprise Development Teams

Enterprise development teams working with software supply chains often rely on an unchecked variety of supply from OSS projects where each developer or development team can make their own sourcing and procurement decisions. The effort of managing 2,778 different projects and 8,200 unique releases (SEE SECTION 4.2.1) can introduce significant drag on development and is contrary to an enterprise’s need to develop faster as part of any agile, continuous delivery or DevOps practice.

Choosing open source projects should be considered an important strategic decision for enterprise software development organizations. Different components demonstrate healthy or poor performance that impacts the overall quality of their releases. Therefore, MTTU should be an important metric when deciding which components to utilize within your software supply chains. Rapid MTTU is associated with lower security risk and is also more accessible from public sources than other metrics enterprises might want to rely upon such as vulnerability data.

Just as traditional manufacturing supply chains intentionally select parts from approved suppliers and rely upon formalized procurement practices — enterprise development teams should adopt similar criteria for their selection of OSS components. This practice ensures the highest quality parts are selected from the best and fewest suppliers — a practice Deming recommended for decades. Implementing selection criteria and update practices will not only improve quality, but can accelerate mean time to repair when suppliers discover new defects or vulnerabilities. Chapter 4 will further explore the impact of OSS component selection on overall application quality.

KEY POINTS

- We recommend projects focus on accelerating and maintaining rapid MTTU.
- Teams should aim for a minimum of four releases annually, and aim to upgrade at least 80% of their dependencies with every release.
CHAPTER 4
Exemplary Dev Teams
Benefits of DevSecOps and Automated Open Source Governance
4.1 The Enterprise Continues to Accelerate

Software innovation has become the last path to differentiation in most competitive industries. Companies must transform their approach to digital innovation or face loss of market share. There is $2 trillion spent annually in software development, but most is labor-oriented and there is a growing need to automate more of the software development process. Thus, exemplary engineering teams are embracing DevOps practices and automated tools to manage third-party dependencies and minimize open source risk.

Forty-seven percent of development teams now deploy to production multiple times a week.21 Furthermore, DORA’s 2018 State of DevOps Report provides strong evidence that organizations adopting DevOps practices are experiencing remarkable results, including:

- DevOps teams deploy code 46x more frequently — meaning they deploy multiple times per day instead of once a week or less.
- DevOps teams have a 7x lower change failure rate — meaning changes to production fail 7.5% of the time instead of 38.5%.
- High performing DevOps teams are 1.75x more likely to extensively use open source software and 1.5 times more likely to expand open source usage in the future.22

In order to survive and thrive in today’s application economy the best development teams are actively embracing open source innovation, dependency management practices, and automated tooling for open source governance.

4.2 Analysis of 12,000 Large Enterprises

This year’s research analyzed the Java open source consumption patterns of 12,000 enterprise development teams to understand average consumption patterns, the diversity of OSS components they relied upon and the number of releases of those projects they consumed. Open source component release download patterns were observed for calendar year 2018. Downloads were observed across enterprises representing 173 countries.

4.2.1 Analysis of Open Source Downloads

**QUESTIONS:** How many component releases are downloaded by companies each year? How many OSS projects and releases are represented?

**FINDINGS:** In 2018, the average enterprise downloaded 313,000 open source component releases — representing an increase of 84% year over year.

On a country-by-country basis, downloads patterns revealed interesting variation. For example, the average organization in Germany downloaded 436,000 component releases, followed by France with 324,000, the United States with 309,000, and the United Kingdom with 248,000 downloads.

For the study population, downloads represented an average of 2,778 open source components, including 8,200 unique component releases. This means the enterprises consumes an average of three releases per component. One the high end, 243 (2%) organizations used over almost five (4.61) release versions. On the low end, 1,470 (12%) organizations averaged below two (1.82) release versions.

**KEY POINTS**

- There is $2 trillion spent annually in software development, but most is labor-oriented and there is a growing need to automate more of the software development process.
- In 2018, the average enterprise downloaded 313,000 open source component releases.
Researchers noted that it is not uncommon to see downloads of 50 – 60 versions of a specific project over the year long observation period.

### 4.2.2 Utilization of Repository Managers

**QUESTION:** How frequently are repository managers used in the download process?

**FINDINGS:** Software developers doubled their use of repository managers in 2018 as they sought more efficient and higher velocity practices for OSS consumption. Over 9 million developers now use repository managers as part of their development tool set. Even as the number of repository manager instances grow, their use as a primary download path within development teams has not reached significant levels. Across the 12,000 organizations analyzed, component release downloads via repository managers totaled 5.3% (compared to 1.8% globally).

By contrast, 305 (3%) of the 12,000 organizations demonstrated high utilization of repository managers for more than 50% of their downloads. This group was utilizing repository managers 14.2 times more on average the other organizations observed.

Repository managers not only accelerate the development process by caching component releases locally, but they also can be used to limit the number of paths releases can travel to make their way into an enterprise. Limiting the number of paths is one of the first steps toward controlling and auditing software supply chain behaviors for an enterprise.

### 4.3 Component Releases Make Up 85% of a Modern Application

Open source components are pervasive in software development today. An analysis of over 500 applications revealed the average application contains over 460 software component releases, of which 85% were open source. In the same study, it was not uncommon to see applications assembled from 2,000 – 4,000 OSS component releases.

In JavaScript development, the number of npm packages per application is even greater. According to npm.org, “the average modern web application has over 1,000 modules, and trees of over 2,000 modules are not uncommon. In fact, 97% of the code in a modern web application comes from packages downloaded from the npm repository. An individual developer is responsible only for the final 3% that makes their application unique and useful.”

### FIG. 4A Percentage of Downloads via Repository Managers

(12K Organizations Analyzed)

<table>
<thead>
<tr>
<th>Organizations</th>
<th>Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>305</td>
<td>50-100%</td>
</tr>
<tr>
<td>582</td>
<td>20-49%</td>
</tr>
<tr>
<td>750</td>
<td>10-19%</td>
</tr>
<tr>
<td>10,363</td>
<td>&lt;10%</td>
</tr>
</tbody>
</table>

**KEY POINTS**

- Across the 12,000 organizations analyzed, component release downloads via repository managers totaled 5.3%
- The Exemplars were utilizing repository managers 14.2 times more on average the other organizations observed.
- An analysis of over 500 applications revealed the average application contains over 460 software component releases, of which 85% were open source.
- 97% of the code in a modern web application comes from packages downloaded from the npm repository.
As development teams strive to deploy new software faster, the practice of assembling open source component releases into the form of an application screams of efficiency. Developers no longer need to code every line from scratch. Developers can download component releases in seconds that deliver new capabilities, built by experts outside of their organizations who make their code freely available to others.

While component use proliferates every development organization today, management of components in these organizations varies considerably. The best organizations follow software supply chain management principles to ensure the best quality component releases are assembled into their applications.

4.4 Characteristics of Exemplary Development Teams

Jeremy Long, founder of the OWASP Dependency Check project once shared, “Good development teams consider out-of-date libraries a code quality issue. They build time into their schedule to upgrade their dependencies. On the other hand, development teams who do not do this regularly are often afraid to break their build.”

When it comes to software development, exemplary teams are more likely to embrace the following patterns and practices when it comes to open source dependency updates.

4.4.1 Dependency Update Behaviors for Developers

In order to better understand the practices surrounding open source components within development, this year’s researchers surveyed 658 developers in April 2019. The developers were asked to describe their organization’s practices as well as to describe their feelings of how those practices impacted their productivity and enjoyment of work.

Researchers performed an analysis of all the respondent answers and three distinct clusters emerged. Broadly speaking, the clusters demonstrated high, medium, and low (29.3%, 48.6%, and 22.0% of respondents, respectively) degrees of reported pain associated with updating dependencies and patching. Researchers then compared the high and low pain clusters to determine what percentage of respondents answered “strongly agree” to the survey questions. Stark differences were found between them (SEE FIGURE 4B).

4.4.1.1 Updating Open Source Dependencies

**QUESTION:** Is updating dependencies scheduled as part of your daily work?

**FINDINGS:** Researchers found that 38% of the developers surveyed updated dependencies as part of their daily work, yet Exemplars were 10x more likely to schedule dependency updates as part of their daily work.

4.4.1.2 Using the Latest Versions of Dependencies

**QUESTION:** Do you strive to use the latest version (or latest-N) of all your dependencies?

**FINDINGS:** From the overall survey population, 46% of developers strove to use the latest version of all of their open source dependencies. Further analysis revealed that Exemplars were 6.2x more likely to use the latest version (or latest-N) of all of their dependencies.

Using the latest versions of dependencies is its own reward. As noted previously in Chapter 3, Exemplar components demonstrated 3.4 times faster MTTR and were 27% more likely to already be protected when new vulnerabilities were discovered.

Then, as will be revealed below in section 4.5, teams using the latest component releases can reduce the presence of vulnerable component releases in their applications by 55%. Therefore, development teams that simultaneously incorporate the latest versions of component releases and procure them from exemplary open source projects can improve the overall quality of their applications.

Using the latest versions of dependencies is its own reward. As noted previously in Chapter 3, Exemplar components demonstrated 3.4 times faster MTTR and were 27% more likely to already be protected when new vulnerabilities were discovered.

**KEY POINTS**

- Jeremy Long, founder of the OWASP Dependency Check project once shared, “Good development teams consider out-of-date libraries a code quality issue. They build time into their schedule to upgrade their dependencies.”
- Exemplars are 10x more likely to schedule dependency updates as part of their daily work.
- Exemplars are 6.2x more likely to use the latest version (or latest-N) of all their dependencies.
- Exemplar components demonstrated 3.4 times faster MTTR and were 27% more likely to already be protected when new vulnerabilities were discovered.
3.2x less likely to consider updating “painful.”
2.6x less likely to consider updating vulnerable component releases “painful.”

Strive to use the latest version (or latest-N) of all dependencies.
**EXEMPLARS:** 6.2x more likely

Schedule update dependencies as part of daily work.
**EXEMPLARS:** 10x more likely

Have a process to proactively remove problematic or unused dependencies.
**EXEMPLARS:** 9.3x more likely

Have automated tools to track, manage, and/or ensure policy compliance of dependencies.
**EXEMPLARS:** 12x more likely

Use some process to add a new dependency (e.g., evaluate, approve, standardize, etc.)
**EXEMPLARS:** 11x more likely

**FIG. 4B Traits of Exemplary Development Teams**
4.4.1.3 USING PROCESSES TO ADD A NEW DEPENDENCY

**QUESTION:** Do you use some process to add a new dependency (e.g., evaluate, approve, standardize, etc.)?

**FINDINGS:** From the overall survey population, 50% of developers said they relied on some process to add a new open source dependency. When examining the Exemplar cluster, researchers found them to be 11x more likely to have a process in place to add a new dependency (e.g., evaluate, approve, standardize, etc.).

4.4.1.4 PROACTIVELY REMOVING DEPENDENCIES

**QUESTION:** Do you have a process to proactively remove problematic or unused dependencies?

**FINDINGS:** When it came to removing problematic component releases (e.g., those with security vulnerabilities), only 30% admitted to having a process in place. Exemplars were 9.3x more likely to proactively remove problematic or unused dependencies.

4.4.1.5 USING AUTOMATION TO MANAGE DEPENDENCIES

**QUESTION:** Do you have automated tools to track, manage, and/or ensure policy compliance of your dependencies?

**FINDINGS:** Use of automated solutions can expedite dependency management. This year’s survey revealed 37% of the overall population relied on automation to manage dependencies. A closer look at Exemplars demonstrated that they were 12x more likely to have automated tools to track, manage, and/or ensure policy compliance of their dependencies.

4.4.1.6 CHARACTERIZING THE EFFORT TO UPDATE DEPENDENCIES

**QUESTION:** Do you consider updating dependencies painful?

**FINDINGS:** Updating dependencies is not a favorite past time of developers. In the overall survey population, 51% of developers agreed that updating was considered painful. The benefit of applying processes and updating to the latest versions of dependencies paid off for the Exemplars. Researchers found Exemplars to be 3.2x less likely to consider updating dependencies to be “painful.”

4.4.1.7 CHARACTERIZING THE EFFORT TO UPDATE VULNERABLE COMPONENTS

**QUESTION:** Do you consider updating vulnerable component releases to be painful?

**FINDINGS:** In the survey, 52% reported the practice of updating vulnerable component releases as “painful.” Again, the benefit of applying processes, updating to the latest N-versions, and regularly removing problematic dependencies paid off for the Exemplars. Exemplars surveyed were 2.6x less likely to consider updating vulnerable components to be “painful.”

4.4.2 Component Management in the Enterprise

This year, our researchers surveyed 5,558 development and DevOps professionals as part of the 2019 DevSecOps Community Survey. Survey participants were asked to self-identify their DevOps maturity level into different categories, where three clusters emerged. Broadly speaking, there were those with high levels of DevOps maturity, those with improving maturity, and those with low maturity or no DevOps practice.

Stark differences emerged between the high maturity (Exemplar) and low maturity (No DevOps) clusters as they responded to questions about open source component controls, policies, and breaches as seen across the following practices.

**KEY POINTS**

- When examining the Exemplar cluster, researchers found them to be 11x more likely to have a process in place to add a new dependency.
- Exemplars were 9.3x more likely to proactively remove problematic or unused dependencies.
- A closer look at Exemplars demonstrated that they were 12x more likely to have automated tools to track, manage, and/or ensure policy compliance of our dependencies.
- Researchers found Exemplars to be 3.2x less likely to consider updating dependencies to be “painful.”
- Exemplars surveyed were 2.6x less likely to consider updating vulnerable components to be “painful.”
4.4.2.1 GENERATION OF A SOFTWARE BILL OF MATERIALS (SBOM)

QUESTION: Does your organization keep a complete SBOM?

FINDINGS: More rigorous control of open source component releases used in development and operations leads to the use of a Software Bill of Materials (SBOM). SBOMs are used to track and trace which open source component releases have been assembled into an application. The survey revealed that only 53% of Exemplars keep a complete software bill of materials.26 When paired with vulnerability data about the components, SBOMs can be used to quickly identify where defective component releases have been used in applications — whether under development or in production, thereby accelerating remediation efforts.

4.4.2.2 CONTROLLING OPEN SOURCE USED IN DEVELOPMENT

QUESTION: How well does your organization control which open source component releases are used in development?

FINDINGS: 2019 saw more organizations investing in controls that start with keeping an inventory of all component releases used. When asked how well their organizations controlled which open source component releases were used in development, 74% of Exemplars in DevOps practices remarked that their organization was “completely locked down” or had “some standards” in place. By contrast, 48% of developers in organizations without a DevOps practice claimed to have “no standards” in place.27

4.4.2.3 EMBRACEMENT OF AUTOMATED TOOLS TO ENFORCE COMPLIANCE AND INFORM DEVELOPERS

QUESTION: Does your organization have an open source policy and do you follow it?

FINDINGS: The survey revealed that 57% of organizations have an open source governance policy in place.28 These organizations rely on policies supported by a mix of formal documentation, OSS governance committees, automated toolsing, and tribal knowledge.

The survey also revealed that automation had a significant impact on whether or not open source policies were followed. In Exemplars DevOps practices where more automated OSS governance solutions have been deployed, 62% of developers

FIG. 4C Adoption of a Software Bill of Materials

Adoption of Open Source Governance Policies

<table>
<thead>
<tr>
<th></th>
<th>Exemplary DevOps Practices</th>
<th>No DevOps Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have a complete SBOM</td>
<td>53%</td>
<td>21%</td>
</tr>
<tr>
<td>Meaningful component controls</td>
<td>47%</td>
<td>79%</td>
</tr>
</tbody>
</table>

SOURCE: 2019 DEVSECOPS COMMUNITY SURVEY (SONATYPE)

KEY POINTS

- The survey revealed that only 53% of Exemplars keep a complete software bill of materials.
- 74% of Exemplars in DevOps practices remarked that standards were in place for controlling use of open source components.
- 57% of organizations have an open source governance policy in place.
remarked that their organization had a policy and that it was followed. By comparison, developers in organizations with little to no DevOps practices, only 25% of their developers were aware of and followed the policy. When automation of OSS policies is present, developer adherence is 150% stronger.\textsuperscript{29}

4.4.2.4 INFORMING DEVELOPERS OF SECURITY RELATED ISSUES

**QUESTION:** How are you informed of InfoSec and AppSec issues?

**FINDINGS:** One of the reasons why adherence is stronger in exemplary organizations is that OSS component attributes (e.g., version numbers, vulnerability information, license descriptions) and policy information are delivered inside developer tooling. For developers in Exemplars DevOps practices, 63% are informed of application security issues from within their own development tools — meaning they don’t have to leave their common tool sets to receive alerts and information.

Developers inside exemplary teams were 62\% more likely to receive notice of application security issues from within their tool sets compared to developers with little to no DevOps practice in place.\textsuperscript{30}

**KEY POINTS**

- When automation of OSS policies is present, developer adherence is 150\% stronger.
- For developers in Exemplars in DevOps practices, 63\% are informed of application security issues from within their own development tools.

---

**FIG. 4D** Age of Components Used in Managed Software Supply Chains
(Analysis of Java Components Across 68,000 Applications)

More than half (51.3\%) of all available components are less than 3 years old.

**FIG. 4E** Percentage of Components with Known Vulnerabilities

Components less than 3 years old have 65\% fewer known vulnerabilities.
4.5 Rewards for Exemplary Development Teams

Development teams who regularly update their open source dependencies and manage their software supply chains can dramatically reduce their risk exposure from the use of vulnerable component releases. For 2019, researchers performed an extensive analysis of open source component releases used in managed and unmanaged supply chains.

The first set of analyses focused on managed software supply chains. The team looked at open source component releases that were used within 85,000 applications. Analysis of the Java open source component releases revealed the latest versions had the lowest percentage of known defects. Component releases under three years in age reflected security defect rates of 9.3% — representing 45% of OSS parts used in the applications.

Researchers also evaluated older component releases used within managed software supply chains. Analysis reveals that component releases over three years in age demonstrated security defect rates at 15.3% — or 65% higher defect rates compared to the newer component. The older component releases represented 55% of the parts used within managed software supply chains.

The second set of analyses focused on unmanaged software supply chains where automated open source governance solutions were not present. The average application assembled within this set revealed a security defect rate of 21%. While researchers were not able to assess the age of component releases across the population, defect rates here were 40 – 100% higher when compared to components used within managed software supply chains.

Managed supply chains make better software. For organizations who tamed their supply chains, the rewards were impressive: use of known vulnerable component releases was reduced by 55%.

**KEY POINTS**

- Component releases under three years in age reflected security defect rates of 9.3%.
- Analysis reveals that component releases over three years in age demonstrated security defect rates at 15.3%.
- Applications in unmanaged software supply chains revealed security defect rates of 21%.
- For organizations who tamed their supply chains, the rewards were impressive: use of known vulnerable component releases was reduced by 55%.
CHAPTER 5
The Changing Landscape
Vulnerabilities, Adversaries, and Government Influence
5.1 Deming Emphasizes Building Quality In

In 1945, W. Edwards Deming started advising Japanese manufacturers to detect and fix defects at the beginning of the manufacturing process. By the 1960s, Deming’s TQM practices were an intrinsic part of the Japanese culture and were playing rise to their global dominance. Now tied into high-performance production processes, six-sigma manufacturing today aims a defect rate goal of 3.4 parts per million.

Following on the lessons of Deming, Jez Humble and Dave Farley, in their seminal book *Continuous Delivery* (2010), advised development teams to “build quality in.” Further echoing those remarks three years later, Gene Kim advised readers to “emphasize performance of the entire system and never pass a defect downstream” as he introduced “the three ways of DevOps” inside *The Phoenix Project*.

These recommendations would come to form the basis for the DevSecOps movement: release faster, build quality into your applications, and automate security controls with minimal friction.

In order to survive and thrive in today’s application economy the best development teams are actively pursuing open source fueled innovations, DevSecOps transformations, and automation-based controls of component releases flowing through their software supply chains.

To better understand how defective and known vulnerable component releases flow through software supply chains, we first have to look at public open source repositories. In general, a public repository is created to serve components for a given development language. For example, Python, Java, JavaScript, Ruby, and Rust components are all served from repositories specific to their development language.

**FIG 5A The Percentage of Vulnerable Java Components Downloaded Over Four Years**

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>6.1%</td>
</tr>
<tr>
<td>2016</td>
<td>5.5%</td>
</tr>
<tr>
<td>2017</td>
<td>12.1%</td>
</tr>
<tr>
<td>2018</td>
<td>10.3%</td>
</tr>
</tbody>
</table>

**KEY POINT**

- Now tied into high-performance production processes, six-sigma manufacturing today aims a defect rate goal of 3.4 parts per million.
5.2 Tracing Vulnerable Component Release Downloads Across Software Supply Chains

Public open source repositories are immutable by design. This means, any given public repository is the home to all good (and bad) versions of open source component releases. When new vulnerabilities are discovered in component releases, open source projects cannot simply remove the defective component release from the repository, and for good reason. In order for developers to remediate a vulnerable component release, they need access to the defective component release. They first rebuild the application to its original state using the defective version, and replace the vulnerable component release with the safe version.

5.2.1 Mutuum Cavete (Borrower Beware)

It is important to understand that no metadata about a component release downloaded from the public repositories is made available to developers by default (e.g., known security vulnerabilities, release age, observed licenses, interoperability changes, etc.).

Without sufficient due diligence by developers on their component release selections, a vulnerable component download behaves exactly like a safe download. Over the past five years that Sonatype has been tracking downloads from the Central Repository for this report, the percentage of vulnerable Java component releases consumed has ebbed and flowed. In 2018, across billions of open source component release downloads, 1 in 10 (10.3%) had known security vulnerabilities.

This represents a slight decrease from 2017’s peak of 12.1% (1 in 8) but when considered as a quality benchmark for managed supply chains it falls far short of expectations. As mentioned earlier, six sigma manufacturing goals aim for defect rates of 3.4 parts per million. At today’s defect rate of 1 in 10 downloads, software component releases procured by development teams now sit at one sigma.

Poor hygiene practices are also well documented within the JavaScript community. In a 2018 npm survey of over 33,000 worldwide developers, 83% expressed concern about whether the open source software they use is secure (up from 77% in 2017), and 58% believe that there aren’t satisfactory methods for evaluating whether code is safe. Furthermore, an October 2018 report from npm revealed that 51% of JavaScript package downloads contained known vulnerabilities. Eleven percent (11%) of the downloaded component releases analyzed were rated critical and 35% demonstrated high vulnerability ratings.

At the next stop along the software supply chain, researchers analyzed downloads to repository managers. Analysis of downloads to Nexus and Artifactory repository managers (i.e., the local warehouses of software supply chains) by development teams reveals that 7.5% were known vulnerable. Once cached in a private, local repository manager, the component release can be served an unlimited number of times to developers within that enterprise.

5.2.2 A Faustian Bargain?

The capabilities that make open source libraries so attractive to development organizations, also make them risky. While component releases are free for developers to download, they are not created equal, and all of them have a cost with respect to maintenance over time.
Our study of 12,000 enterprise software development organizations revealed average annual component release downloads of 313,000. Further analysis of those downloads reveals that 27,704 (8.8%) included at least one known security vulnerabilities. Just as well, not all security vulnerabilities are created equal. Of the 27,704 vulnerable downloads, 67% had Common Vulnerability Scoring System (CVSS) at 7.0 or above on a 10 point scale. Thirty percent (30%) had CVSS scores above 9.0 on a 10 point scale.

Minor fluctuations in the percentage of vulnerable download were seen on a country by country basis: United States (8.9%), France (8.8%), United Kingdom (8.8%), and Germany (8.1%).

5.3 Adversaries Increasingly Target Open Source Components

The 2019 DevSecOps Community Survey of 5,558 development professionals, highlighted a 71% increase in confirmed or suspected open source related breaches since 2014 (the same year the notorious OpenSSL Heartbleed vulnerability). Minor fluctuations in the percentage of vulnerable download were seen on a country by country basis: United States (8.9%), France (8.8%), United Kingdom (8.8%), and Germany (8.1%).

The percentage of open source related breaches dropped 7% compared to the 2018 survey responses. The slight decline in breaches between the two surveys may be attributed to improved open source hygiene and investments made by some organizations following the Equifax breach.

KEY POINTS

- Further analysis of downloads in one study revealed that 8.8% included at least one known security vulnerabilities.
- Minor fluctuations in the percentage of vulnerable download were seen on a country by country basis: United States (8.9%), France (8.8%), United Kingdom (8.8%), and Germany (8.1%).
- The 2019 DevSecOps Community Survey highlighted a 71% increase in confirmed or suspected open source related breaches since 2014.
— made public in late 2017. Still, with 1 in 4 survey participants in this year’s reporting breaches in the past 12 months — breaches remain at epidemic levels.35

5.3.1 A Post-Equifax Look at Apache Struts Vulnerabilities

According to Fortinet’s Q4’2018 Threat Report, vulnerable instances of Apache Struts remain the most prevalent exploit. “Demonstrating that the internet never forgets, the Apache Struts exploit (associated with CVE-2017-5638) has been a top detection since its role in the infamous Equifax breach back in 2017. More recently, attackers have been using this exploit as a way to implement crypto-jacking functions on compromised machines.”36

Even more alarming is the trend in elective downloads of vulnerable Struts component releases. According to Sonatype’s analysis of Struts downloads from the Central Repository, the volume of monthly vulnerable downloads continued to increase following its link to the breach at Equifax.

KEY POINT

One year after the breach announcement, monthly vulnerable Struts downloads had increased 11% to 2.1 million.
One year after the breach announcement, monthly vulnerable Struts downloads had increased 11% to 2.1 million.

5.3.2 Event-Stream: Malicious Code Injection Targeting Cryptocurrency

In the 2018 State of the Software Supply Chain Report, we advised that adversaries had compressed the time between vulnerability announcement and exploit from 45 to 3 days. In some instances, by injecting malicious code into open source projects on the supply side, time to exploit was further reduced to zero.

Such was the case with the socially engineered code injection for the JavaScript npm package known as event-stream in November 2018. The injection of malicious code into event-stream was accomplished when its developer unwittingly handed his credentials to an adversary who offered to take over maintenance responsibilities. The npm package is downloaded 2 million times per week.

Further investigation of this exploit determined it was targeted as a CoPay’s Bitcoin wallet and “designed to harvest account details and private keys from accounts having a balance of more than 100 Bitcoin or 1000 Bitcoin Cash.”

5.3.3 Bootstrap-Sass: Malicious Injection of a Back Door

Bootstrap-sass is an open source framework that enables web designers to quickly build a site. With over 28 million downloads through March 2019, it is extremely popular among developers.

On March 27 of this year, Derek Barnes, a software developer whose code relied on the popular RubyGems bootstrap-sass component had a build fail. Suspicious, Derek decided to do some research and noticed that “someone” had removed a version of the library (Bootstrap-Sass v3.2.0.2) and immediately released a new version, moments later, v3.2.0.3. He was suspicious why “someone” would modify the library on RubyGems — but not in GitHub, where the library’s source code is managed. What Barnes uncovered was sobering: another attack on open source and the software supply chain that underpins so much of modern innovation.

Alarmed by the abnormal release, he alerted the project which led to a backdoor being discovered in the code. The malicious component version (3.2.0.3) was then removed from the RubyGems repository — and the Bootstrap-Sass team revoked access to RubyGems for the developer whose account they believed was compromised and used to push the malicious code.

Before being noticed, the vulnerable version of bootstrap-sass was downloaded over 1,400 times.

5.3.4 Agama Malicious Code Injection

In June 2019, the npm, Inc. security team, in collaboration with Komodo, helped protect over $13 million USD in cryptocurrency assets by finding and responding to a malicious code injection vulnerability targeting the users of a cryptocurrency wallet called Agama. The attack focused on getting a malicious package into the build chain for Agama and stealing the wallet seeds and other login passphrases used within the application. According to npm, “the attack was carried out by using a pattern that is becoming more and more popular; publishing a ‘useful’ package (electron-native-notify) to npm, waiting until it was in use by the target, and then updating it to include a malicious payload.”

Over the past two years, a dangerous new trend has emerged. Adversaries are taking advantage of a new attack vector where they are directly injecting vulnerabilities into open source project releases and container images.

5.3.5 Two Years of Malicious Code Injection

The mechanics to the bootstrap, event stream, and electron-native-notify attacks are tricky and yet effective.

While limited in scope, the potential impact of similar attacks through a very popular component can be far-reaching and immediate. When malicious code is injected into software supply chains, adversaries can attack immediately after the code is deployed.

Over the past two years, a dangerous new trend has emerged. Specifically, a series of 16 events triangulate a serious escalation of software supply chain attacks. Adversaries are taking advantage of a new attack vector where they are directly injecting vulnerabilities into open source project releases and container images (see FIGURE 5E).
npm credentials published online.
Affects access to 14% of the npm repo (79,000 packages).

Malicious npm package typosquat.
40 packages harvested over two weeks, collecting credentials used to publish to the npm repository itself.

docker123321 images created on Docker Hub.
Later accused of poisoning a Kubernetes honeypot (Jan 2018), and equated to a crypto-mining botnet (May 2018).

“I’m harvesting credit card numbers and passwords from your site. Here’s how.”
David Gilbertson writes a fictional tale on his blog about creating a malicious npm package.

npm credentials intentionally compromised.
A malicious version of a package from a core contributor to the conventional-changelog ecosystem is published. The package was installed 28,000 times in 35 hours and executed a Monero crypto miner.

Linux distro hacked on GitHub.
Unknown individuals gain control of the Github Gentoo organization, and modified the content of repositories as well as pages within. All code considered compromised.

Homebrew repository compromised.
Accessed in under 30 minutes through an exposed GitHub API token.

Back-doored npm package discovered.
Malicious package injected into event-stream, a popular npm package.
The injected code targets the Copay application and was designed to harvest account details and private keys from accounts having a balance of more than 100 Bitcoin or 1,000 Bitcoin Cash.

Back-doored Gems
bootstrap-sass RCE package discovered.
A malicious version of the popular bootstrap-sass package, downloaded a total of 28 million times to date, and with 1.6K dependencies, is published to the RubyGems repository.

Deleted go-bindata account resurrected by an unknown user.
After a developer deleted their GitHub account, someone immediately grabbed the ID — inheriting the karma instilled in that id and calling into question packages and sources.

Back-doored npm package discovered.
npm security team responds to reports of a malicious back door in the get-cookies module, published in March. Despite being deprecated, mailparser still receives about 64,000 weekly downloads.

Compromised JavaScript package caught stealing npm credentials.
A hacker gains access to a developer’s npm account and injects malicious code into a popular JavaScript library called eslint-scope, a sub-module of the more famous ESLint, a JavaScript code analysis toolkit.

Malicious package injected into event-stream, a popular npm package.
The injected code targets users of a cryptocurrency wallet called Agama, focusing on getting into the build chain and stealing the wallet seeds and other login passphrases used within the application.

Cryptocurrency attack via malicious code injection.
Malicious code targets users of a cryptocurrency wallet called Agama, focusing on getting into the build chain and stealing the wallet seeds and other login passphrases used within the application.
5.4 Government and Industry Apply New Standards to Secure Software Development

While Exemplars are reducing the use of known vulnerable open source component releases in the development application, this does not exonerate them from duties to manage components over time. Secure software practices extend from early development through the active life of an application in the market.

With an ever increasing number of application breaches occurring, standards bodies and government are stepping in to hold development organizations accountable for the quality and security of code they assemble and develop.

5.4.1 New PCI Secure Software Development Standards

In January 2019, the Payment Card Industry Security Standards Council introduced important new standards for software development:

- PCI Secure Software Standards,\(^43\) and
- PCI Secure Software Lifecycle (Secure SLC) Standard\(^44\)

The new standard requires organizations to govern their use of open source software, and it states that any application utilized as part of the payment process, must be secure by design. Specifically, when it comes to the use of open source components and third-party libraries organizations are responsible for ensuring they, as well as any of their vendors, have:

- An up to date inventory of open-source component releases utilized in the software
- A process for identifying known vulnerabilities within open source component releases
- 360 degree monitoring of open source component releases throughout the SDLC
- A policy and process to immediately remediate vulnerabilities as they become known.

To effectively utilize open source components at scale, the PCI standard also advises organizations to generate a software bill of materials (SBOM) so they can easily track and trace the location of every single component release embedded within their production software applications.

5.4.2 National Telecommunications and Information Administration’s (NTIA) SBOM Initiative

The Commerce Department’s National Telecommunications and Information Administration (NTIA) is considering requiring companies to list their sources of software parts to protect the U.S. software supply chains.

According to the NTIA, their “cybersecurity multistakeholder process will focus on Software Component Transparency. Participants will explore how manufacturers and vendors can communicate useful and actionable information about the third-party software components that comprise modern software and IoT devices, and how this data can be used by enterprises to foster better security decisions and practices.”\(^45\)

The NTIA effort comes as the government agencies grow more suspicious of known vulnerable software components being used in application development and increasing awareness of malicious code injection attacks. The NTIA initiative has broad support across Federal agencies and the private sector, as they work together to define standards around a software bill of materials (SBOM).

---

**KEY POINTS**

- The SBOM permits organizations to make informed risk decisions about which technologies to purchase and use based on known vulnerability information.
- When new vulnerabilities are discovered, an SBOM allows organizations to quickly identify their exposure and to take appropriate steps in response.”
5.4.4 U.S. House Energy and Commerce Committee

In December 2018, the U.S. House Energy and Commerce Committee released its Cybersecurity Strategy Report. The report details the importance and priority for utilizing Software Bill of Materials (SBOM) for minimizing supply chain risks related to the use of open source software components in modern application development.46

The report states, the “SBOM becomes an ingredients list for a given piece of technology, listing the hardware, software, and other relevant components that it contains or relies upon. This creates two primary benefits. First, it permits organizations to make informed risk decisions about which technologies to purchase and use based on known vulnerability information. Second, when new vulnerabilities are discovered, it allows organizations to quickly identify their exposure and to take appropriate steps in response.”

Next the authors from the Energy and Commerce Committee pointed out that it “was not that organizations did not know which software was vulnerable... it was that they did not know which pieces of technology that they depended on included it. The SBOM minimizes the number of unknown unknowns with which organizations must contend, and greatly increases their ability to protect themselves, their users, and ultimately society.”

5.4.5 U.S. Food and Drug Administration

In October 2018, the FDA released guidance for cybersecurity management of medical devices. Similar to recommendations delivered by the House Energy and Commerce Committee, the FDA’s report called for a Cybersecurity Bill of Materials (CBOM).

The aim of the CBOM is to provide a list of “commercial, open source, and off-the-shelf software and hardware components to enable device users (including patients, providers, and healthcare delivery organizations (HDOs) to effectively manage their assets, to understand the potential impact of identified vulnerabilities to the device (and the connected system), and to deploy countermeasures to maintain the device’s essential performance.”

The FDA report also suggested that MDMs may be subject to legal liabilities tied to the distribution of a medical device with a known vulnerability. The report offered, “If a medical device cybersecurity vulnerability allegedly causes bodily harm, the victim may bring product liability claims against the MDM.” The FDA warned that a plaintiff’s claim could be based on the “failure to provide appropriate warnings about the risk of a cybersecurity vulnerability, or failure to appropriately monitor for the existence of vulnerabilities and appropriately address them once identified.”47

KEY POINTS

- In October 2018, the FDA released guidance for cybersecurity management of medical devices. The FDA’s report called for a Cybersecurity Bill of Materials (CBOM).
- The FDA report suggests that MDMs may be subject to legal liabilities tied to the distribution of a medical device with a known vulnerability.
Conclusion

Decades ago, W. Edwards Deming taught key principles to significantly improve the effectiveness and quality of business manufacturing processes. Deming deftly advocated selecting the best suppliers and emphasized continuous improvement. He advised eliminating the need for inspection on a mass basis by building quality into products.48

Businesses racing to deliver better value to their customers — and differentiate from competitors — are embracing Deming’s principles within their open source based software development practices. Software development is evolving from artisan based creations to practices that more closely resemble high velocity parts assembly. This is reflected in the exponential growth of supply and demand for open source components. We’ve observed double and triple digit growth in open source component ecosystems for a decade, and there is no slowdown in sight.

The purpose of this report was to share with you what we observed across software supply chains. Our findings are clear. Velocity does not have to come at the cost of reduced security.

Exemplary open source project initiatives benefit tremendously from higher code commit and release frequencies. They also do an outstanding job of managing their dependencies. At the same time, Exemplars in enterprise are benefiting from processes that support using the latest component versions. They also embrace automated practices to reduce the presence of known vulnerabilities.

Our deep examination of consumption patterns, development practices, and cybersecurity hygiene revealed:

- 18x faster median time to update dependencies for exemplary open source components
- 6.2x more likely for exemplary enterprise development teams to use the latest open source component version (or latest-N)
- 12x higher use of automation to manage open source dependencies in exemplary enterprise development teams
- 55% reduction in the use of vulnerable open source components within managed software supply chains

Management of software supply chains is not simply ensuring quality at velocity. Our supply chains are being attacked by adversaries in new and creative ways. The result: open source related breaches have jumped 71% over the past five years.

As enterprises look for guidance to improve their software development and security practices, industry standards groups are introducing open source awareness policies. Simultaneously, we’ve seen government agencies racing to employ new policies and legislation to protect citizens, businesses, and critical infrastructure.

One thing is clear, exemplary software development practices that deliver high quality, improved security at high velocity are not rare. They are being employed today in large numbers and serve as benchmarks for others to strive for and achieve.

Thank you for reading our 5th annual State of the Software Supply Chain Report. We hope you found it useful. And, we welcome your feedback.
Sources

1 Sonatype, npmjs.org, pypi.org, NuGet.org
2 hub.docker.com
3 Sonatype and modulecounts.com
4 npmjs.org and modulecounts.com
5 developers.slashdot.org/story/17/01/14/0222245/nodejs-npm-is-now-the-largest-package-registry-in-the-world
6 www.idc.com/getdoc.jsp?containerId=US44363318
7 www.oracle.com/java/java9.html
8 twitter.com/seldo/status/1105987692305604608
9 9.7M developers use JavaScript: appdevelopermagazine.com/
10 Sonatype and rubygems.org
11 Sonatype and nuget.org
12 https://chaoss.community
13 https://libraries.io
14 Sonatype.com/nexus-intelligence
15 The (Application) Patching Manifesto (Jeremy Long): www.youtube.com/watch?time_continue=144&v=qVZgTRJ290
16 Why and How Java Developers Break APIs (Aline Brito, Laerte Xavier, Dr. Andre Hora, Dr. Marco Tulio Valente): arxiv.org/pdf/1801.05198.pdf
17 A Pearson correlation is a number between -1 and 1 that indicates the extent to which two variables are linearly related. The Pearson correlation is also known as the “product moment correlation coefficient” (PMCC) or simply “correlation.” en.wikipedia.org/wiki/Pearson_correlation_coefficient
18 Kendall’s Tau is a non-parametric measure of relationships between columns of ranked data. The Tau correlation coefficient returns a value of 0 to 1, where: 0 is no relationship, and 1 is a perfect relationship. www.statisticshowto.dasciencemencentral.com/kendalls-tau
20 Sonatype 2019 DevSecOps Community Survey
22 Sonatype and jfrog.com/about/press/jfrog-securer-165-million-investment-to-lead-universal-devops-in-the-enterprise
24 The (Application) Patching Manifesto (Jeremy Long): www.youtube.com/watch?time_continue=144&v=qVZgTRJ290
25 Sonatype 2019 DevSecOps Community Survey
26 Sonatype 2019 DevSecOps Community Survey
27 Sonatype 2019 DevSecOps Community Survey
28 Sonatype 2019 DevSecOps Community Survey
29 Sonatype 2019 DevSecOps Community Survey
30 Sonatype 2019 DevSecOps Community Survey
32 NTIA Software Component Transparency: www.ntia.doc.gov/SoftwareTransparency
34 NTIA Software Component Transparency: www.ntia.doc.gov/SoftwareTransparency
37 Fortinet Threat Report, Q4'2018
38 Details about the event-stream incident: blog.npmjs.org/post/18057383195/details-about-the-event-stream-incident
39 Github.com/urbit/urbit-wallet-generator/issues/4
40 Github.com/twbs/bootstrap-sass
41 Github.com/twbs/bootstrap-sass
42 Plot to steal cryptocurrency foiled by the npm security team: blog.npmjs.org/post/185397814280/plot-to-steal-cryptocurrency-foiled-by-the-npm
45 NTIA Software Component Transparency: www.ntia.doc.gov/SoftwareTransparency
48 Dr. Deming’s 14 Points for Management: deming.org/explore/fourteen-points?apartner=aarp
Appendix A

Acknowledgments
Each year, the State of the Software Supply Chain report is produced to shed light on the patterns and practices associated with open source software development. We began collecting data for our 2019 report from the moment our 2018 report was published.

The report is made possible thanks to a tremendous effort put forth by many team members at Sonatype, including: Derek Weeks, Matt Howard, Joel Orlina, Bruce Mayhew, Gazi Mahmud, Dariush Griffin, Mike Hansen, Brian Fox, Ilkka Turunen, Elissa Walters, Daniel Sauble, Adam Cazzolla, Alex Metry, Andrew Stein, Ken Duck, Kevin Hayen, Kevin Witten, Shade Solon, Alvin Gunkel, and Aaron Massey.

We would also like to offer thanks for contributions big and small from: Hasan Yasar (Carnegie Mellon University Software Engineering Institute), Laurie Voss (npm), Brian Dawson (CloudBees), DJ Schleen (Aetna CVS), Nichole Schimanski (Galois) and Eric Davis (Galois), James Wickett and others across the DevOps and open source development community.

A very special thanks goes out to Melissa Schmidt who created the incredible design for this year’s report.

Finally, we could not have produced this report without the amazing contributions and countless hours of deep analysis from our research partners Gene Kim from IT Revolution and Dr. Stephen Magill, Principal Scientist at Galois & CEO of MuseDev.

About the Analysis
The authors have taken great care to present statistically significant sample sizes with regard to component versions, downloads, vulnerability counts, and other data surfaced in this year’s report. Specifically to Chapter 3, all reported differences are statistically significant ($p < 0.05$) according to a Mann–Whitney U test. While Sonatype has direct access to primary data for Java, JavaScript, Python, .NET and other component formats, we also reference third-party data sources as documented.
Appendix B

23 Metrics Associated with Open Source Project Health

This year’s analysis includes datasets representing development velocity, team size, continuous integration (CI) usage, known security vulnerabilities, component popularity, and other attributes. The objective of this research was to look for top performing projects and characterize their dimensions of excellence. The research also suggests factors that architects, developers, and organizations should consider when making choices about which open source components to use. Our analysis also took into account over 23 metrics associated with open source project health including:

1. **Component_key**: The Maven Group ID + Artifact ID, e.g. org.springframework:spring-core
2. **Median_ttud_median (lower is better)**: The median time to upgrade a dependency. For each dependency D, start the clock when an update to some new version V of the dependency is released. Stop it when the component has a release where the version of D is >= V.
3. **Median_ttu_sec_rel_median (lower is better)**: The median TTU for updates that had a known vulnerability against them at the time the new version was released. This is the best TTR data we have.
4. **Median_ttr_median (lower is better)**: A version of TTR that accounts for vulnerabilities discovered long after a component is released. Clock starts when a vulnerability is reported against a dependency of a component. Clock stops when that component updates the dependency.
5. **Already_protected_percentage (higher is better)**: Records the percentage of security vulnerabilities that don’t apply to this component when they come out because the component was up-to-date with dependencies.
6. **Max_dependency_count**: The number of dependencies of the component.
7. **Avg_not_adopted_avg (lower is better)**: The number of dependency updates that were never applied.
8. **Avg_stale_avg (lower is better)**: The average percentage of component dependencies that are out of date when a new release of that component comes out.
9. **Avg_period_when_current (lower is better)**: Average time each release spends as the “current” release. Reciprocal of release frequency.
10. **Avg_stale_time_avg (lower is better)**: Average period when at least one dependency is out of date. Can probably drop this one.
11. **Stale_time_proportion (lower is better)**: The percentage of time that a component spends being out of date. Can probably drop this too.
12. **Scm_url**: The URL where the code lives.
13. **Two_dots**: The first two components of the group ID.
14. **Three_dots**: The first three components of the group ID.
15. **Central_popularity**: The average number of downloads per day from The Central Repository.
16. **NexusIQ_popularity**: The average number of IQ scans per day.
17. **Stars_count**: Number of stars.
18. **Forks_count**: Number of forks.
19. **Ci-found**: Whether we found a CI script.
20. **Avg-commits-per-month**: Average number of commits per month.
21. **Avg-commits-uniq-devs**: Average number of unique developers committing each month.
22. **Is_commercially_supported**: True if group ID starts with “com”
23. **Is_foundation_supported**: True if there are several projects managed by the same Group ID.
Appendix B

Cumulative Histogram of MTTR

For developers wanting to use secure components, having faster MTTR is desirable. The figures below show a simple histogram of MTTR and MTTU across all components.

Frequency of Median Time to Remediate Vulnerability

Frequency of Median Time to Update Vulnerable Dependencies
Appendix C

Measures of Different Behavior Clusters

### Averages

<table>
<thead>
<tr>
<th>Feature</th>
<th>None (8142)</th>
<th>Small Exemplars (606)</th>
<th>Large Exemplars (595)</th>
<th>Laggards (521)</th>
<th>Features First (280)</th>
<th>Cautious (429)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTU (Days)</td>
<td>245.22</td>
<td>18.74</td>
<td>10.14</td>
<td>692.82</td>
<td>602.94</td>
<td>103.44</td>
</tr>
<tr>
<td>Stale Dependency Percentage</td>
<td>0.64</td>
<td>0.11</td>
<td>0.09</td>
<td>0.98</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>Central Popularity (Average Daily Downloads)</td>
<td>425.5</td>
<td>803.67</td>
<td>4681.28</td>
<td>62.4</td>
<td>297.65</td>
<td>1349.16</td>
</tr>
<tr>
<td>Stars</td>
<td>907.31</td>
<td>191.07</td>
<td>1150.31</td>
<td>452.92</td>
<td>1289.35</td>
<td>1009.85</td>
</tr>
<tr>
<td>Forks</td>
<td>343.84</td>
<td>41.78</td>
<td>651.37</td>
<td>171.51</td>
<td>434.42</td>
<td>351.81</td>
</tr>
<tr>
<td>Number of Dependencies</td>
<td>3.91</td>
<td>3.12</td>
<td>6.14</td>
<td>3.11</td>
<td>3.8</td>
<td>4.11</td>
</tr>
<tr>
<td>Percent Using CI</td>
<td>0.73</td>
<td>0.57</td>
<td>0.69</td>
<td>0.69</td>
<td>0.71</td>
<td>0.78</td>
</tr>
<tr>
<td>Commit Frequency</td>
<td>42.16</td>
<td>19.51</td>
<td>69.34</td>
<td>24.43</td>
<td>56.47</td>
<td>51.25</td>
</tr>
<tr>
<td>Number of Developers (Unique Devs Commiting Per Month)</td>
<td>3.89</td>
<td>1.64</td>
<td>8.91</td>
<td>2.87</td>
<td>5.19</td>
<td>4.67</td>
</tr>
<tr>
<td>Average Period Between New Versions (Days)</td>
<td>128.43</td>
<td>87.62</td>
<td>45.64</td>
<td>244.25</td>
<td>49.99</td>
<td>72.14</td>
</tr>
<tr>
<td>Percent Commercially Supported</td>
<td>0.3</td>
<td>0.25</td>
<td>0.06</td>
<td>0.37</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td>Percent Foundation Supported</td>
<td>0.59</td>
<td>0.57</td>
<td>0.91</td>
<td>0.55</td>
<td>0.65</td>
<td>0.6</td>
</tr>
</tbody>
</table>

### Average Multiples

(as compared to “None” class)

<table>
<thead>
<tr>
<th>Feature</th>
<th>None (8142)</th>
<th>Small Exemplars (606)</th>
<th>Large Exemplars (595)</th>
<th>Laggards (521)</th>
<th>Features First (280)</th>
<th>Cautious (429)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTU</td>
<td>1</td>
<td>0.08</td>
<td>0.04</td>
<td>2.82</td>
<td>2.45</td>
<td>0.42</td>
</tr>
<tr>
<td>Stale Dependency Percentage</td>
<td>1</td>
<td>0.17</td>
<td>0.14</td>
<td>1.55</td>
<td>1.52</td>
<td>1.53</td>
</tr>
<tr>
<td>Central Popularity (Average Daily Downloads)</td>
<td>1</td>
<td>–</td>
<td>11</td>
<td>0.15</td>
<td>0.7</td>
<td>–</td>
</tr>
<tr>
<td>Stars</td>
<td>1</td>
<td>0.21</td>
<td>1.27</td>
<td>0.5</td>
<td>1.42</td>
<td>–</td>
</tr>
<tr>
<td>Forks</td>
<td>1</td>
<td>0.12</td>
<td>1.89</td>
<td>0.5</td>
<td>1.26</td>
<td>–</td>
</tr>
<tr>
<td>Number of Dependencies</td>
<td>1</td>
<td>0.8</td>
<td>1.57</td>
<td>0.8</td>
<td>–</td>
<td>1.05</td>
</tr>
<tr>
<td>Percent Using CI</td>
<td>1</td>
<td>0.78</td>
<td>0.94</td>
<td>0.94</td>
<td>–</td>
<td>1.06</td>
</tr>
<tr>
<td>Commit Frequency</td>
<td>1</td>
<td>0.46</td>
<td>1.64</td>
<td>0.58</td>
<td>1.34</td>
<td>1.22</td>
</tr>
<tr>
<td>Number of Developers (Unique Devs Commiting Per Month)</td>
<td>1</td>
<td>0.42</td>
<td>2.29</td>
<td>0.74</td>
<td>1.34</td>
<td>1.2</td>
</tr>
<tr>
<td>Average Period Between New Versions (Days)</td>
<td>1</td>
<td>0.68</td>
<td>0.35</td>
<td>1.88</td>
<td>0.39</td>
<td>0.57</td>
</tr>
<tr>
<td>Percent Commercially Supported</td>
<td>1</td>
<td>0.85</td>
<td>0.19</td>
<td>1.24</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Percent Foundation Supported</td>
<td>1</td>
<td>–</td>
<td>1.54</td>
<td>0.92</td>
<td>11</td>
<td>–</td>
</tr>
</tbody>
</table>

A dash (–) indicates that the observed difference between groups was not statistically significant (p > 0.05 with a Mann–Whitney U test).
FIG. 3I The Exemplars: Components Demonstrating the Fastest MTTU and Lowest Stale Dependency Counts

Appendix D

- com.github.advantageous:qbit-test-support
- com.github.advantageous:qbit-admin
- com.github.advantageous:qbit-core
- com.github.advantageous:qbit-servlet
- com.github.jakakura:gradle-robovm-plugin
- com.github.kenteye:sdj
- com.github.kzwang:elasticsearch-river-dynamodb
- com.github.kzwang:elasticsearch-transport-redis
- com.github.kzwang:elasticsearch-osem
- com.github.kzwang:elasticsearch-repository-grids
- com.github.mhshams:core
- com.github.michaelrocco:swift2-api-publisher-plugin
- com.github.mitaenge:dozer
- com.github.nikzawa:engine.io-client
- com.github.nwic:contracts
- com.github.oscercamell-cassandra
- com.github.pengrjad:java-telegram-bot-api
- com.github.persapiens:jsf-undertow-boot-starter
- com.github.persapiens:jsf-undertow-bootsfaces-spring-boot-starter
- com.github.persapiens:jsf-jetty-bootsfaces-spring-boot-starter
- com.github.pwittchen:reactiveeventnetwork
- com.github.pwittchen:reactivebeacons
- com.github.ratecommends:gdx-utils
- com.github.richard-ballard:arbee-test-utils
- com.github.richard-ballard:arbee-utils
- com.github.salomonbrys:kodein
- com.github.salomonbrys:kodein-android
- com.github.salomonbrys:kodein-android
- com.github.salomonbrys:kodein-core
- com.github.salomonbrys:kodein-development
- com.github.scala-blitz:scala-blitz
- com.github.salomonbrys:kodein
- com.github.ratrecommends:gdx-utils
- com.github.persapiens:jsf-undertow-boot-starter
- com.github.persapiens:jsf-jetty-bootsfaces-spring-boot-starter
- com.github.pwittchen:reactiveeventnetwork
- com.github.pwittchen:reactivebeacons
- com.github.ratecommends:gdx-utils
- com.github.richard-ballard:arbee-test-utils
- com.github.richard-ballard:arbee-utils
- com.github.salomonbrys:kodein
- com.github.salomonbrys:kodein-development
- com.github.scala-blitz:scala-blitz
- com.github.bladecoder:engine:blade-engine-spinep-plugin
- com.github.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
- com.bladecoder:engine:blade-engine-spinep-plugin
- com.bladecoder:engine:blade-engine
More than 10 million software developers rely on Sonatype to innovate faster while mitigating security risks inherent in open source. Sonatype’s Nexus platform combines in-depth component intelligence with real-time remediation guidance to automate and scale open source governance across every stage of the modern DevOps pipeline. Sonatype is privately held with investments from TPG, Goldman Sachs, Accel Partners, and Hummer Winblad Venture Partners.

Visit www.sonatype.com to learn more.

Sonatype Inc.
Copyright ©2019 – present, all rights reserved. Sonatype and Sonatype Nexus are trademarks of Sonatype, Inc. All other trademarks are the property of their respective owners.

Headquarters
8161 Maple Lawn Blvd, Suite 250
Fulton, MD 20759
United States • 1.877.866.2836

Virginia Office
8281 Greensboro Dr Suite 630
McLean, VA 22102

APAC Office
5 Martin Place, Level 14
Sydney 2000, NSW
Australia

European Office
199 Bishopsgate
London EC2M 3TY
United Kingdom