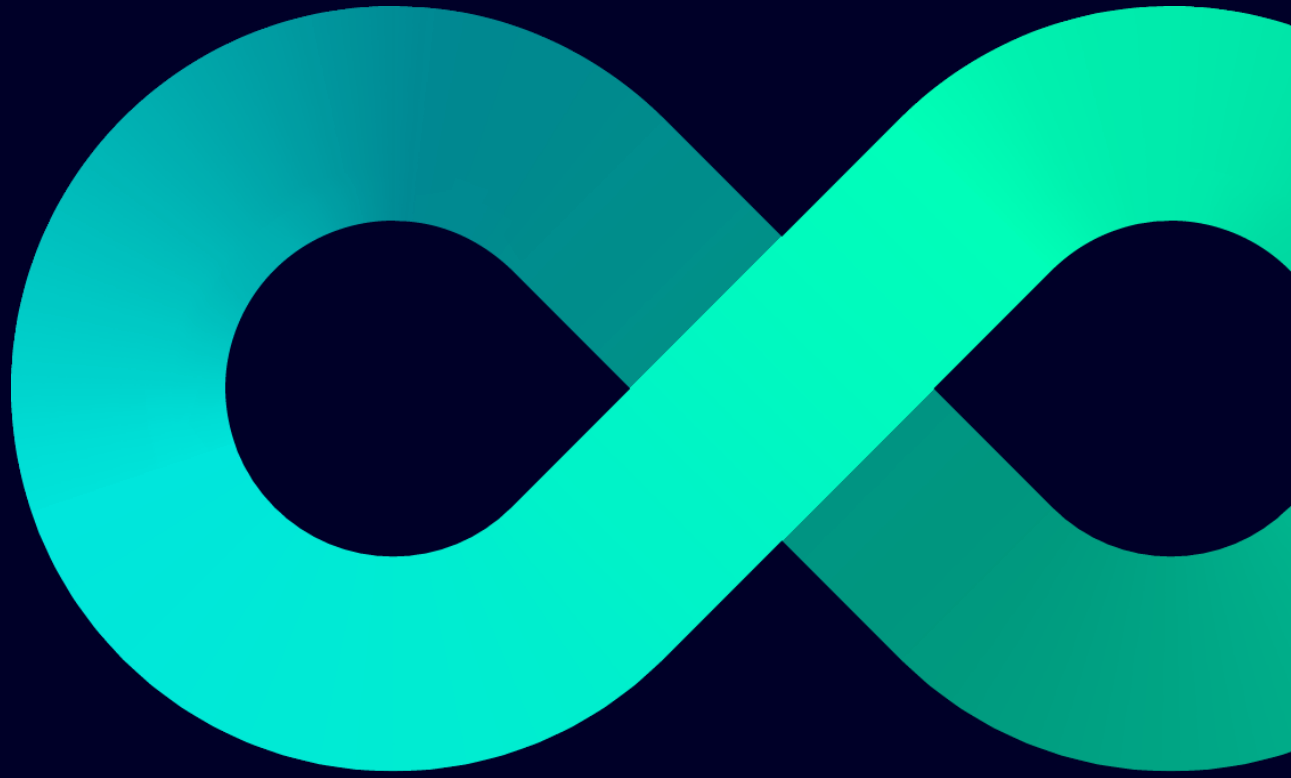
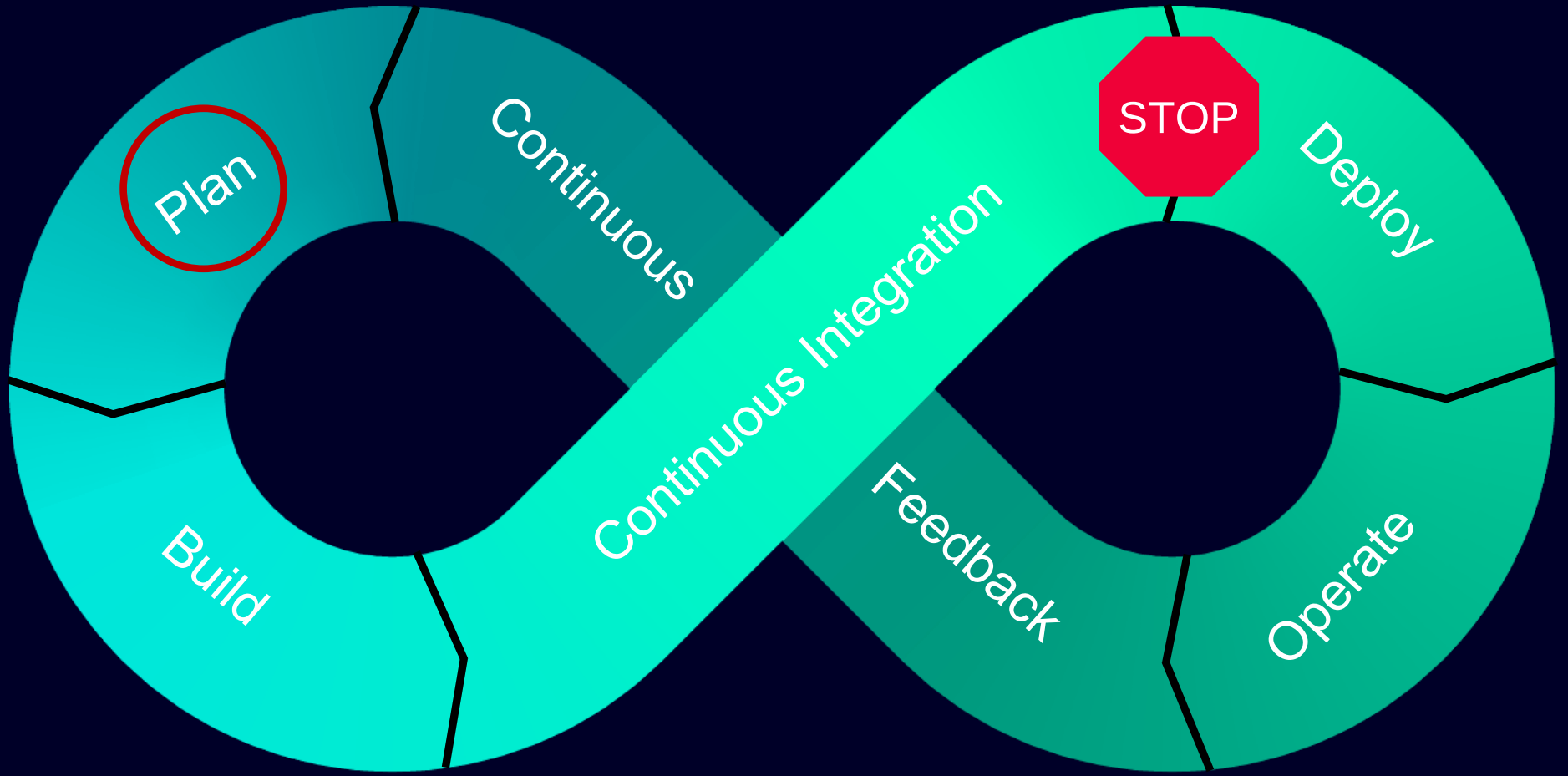
License Compliance und Automatisierung



Bitkom Forum Open Source 2023

SIEMENS

CI/CD – Die Idee alles zu Automatisieren



SIEMENS

Automate **Everything!**

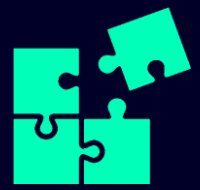
SIEMENS

Siemens

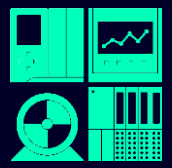
310k Mitarbeiter

47k Entwickler

50+ R&D Standorte



120k Komponenten

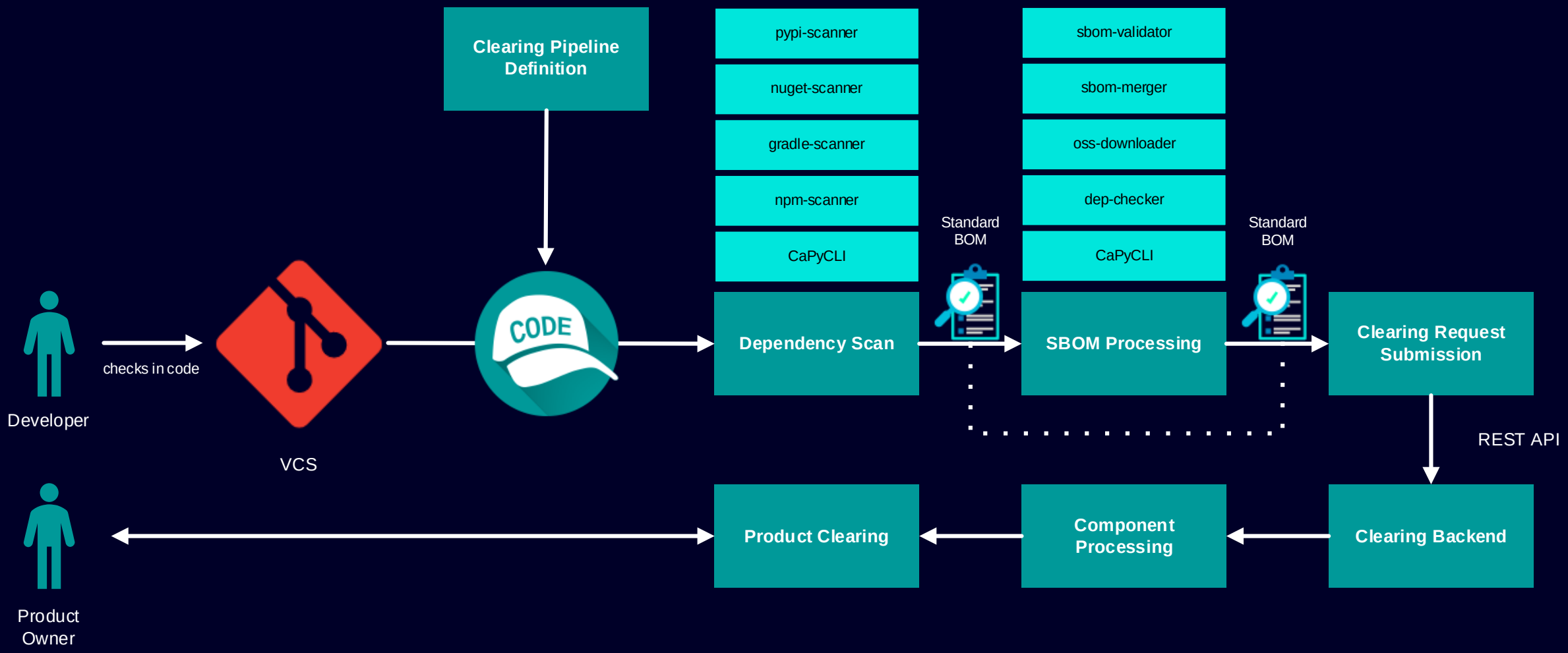


20+ Software Eco-Systems (JavaScript, Java, C#, Go, Python, Lua, Swift, …)

+25k Neue Komponenten pro Jahr

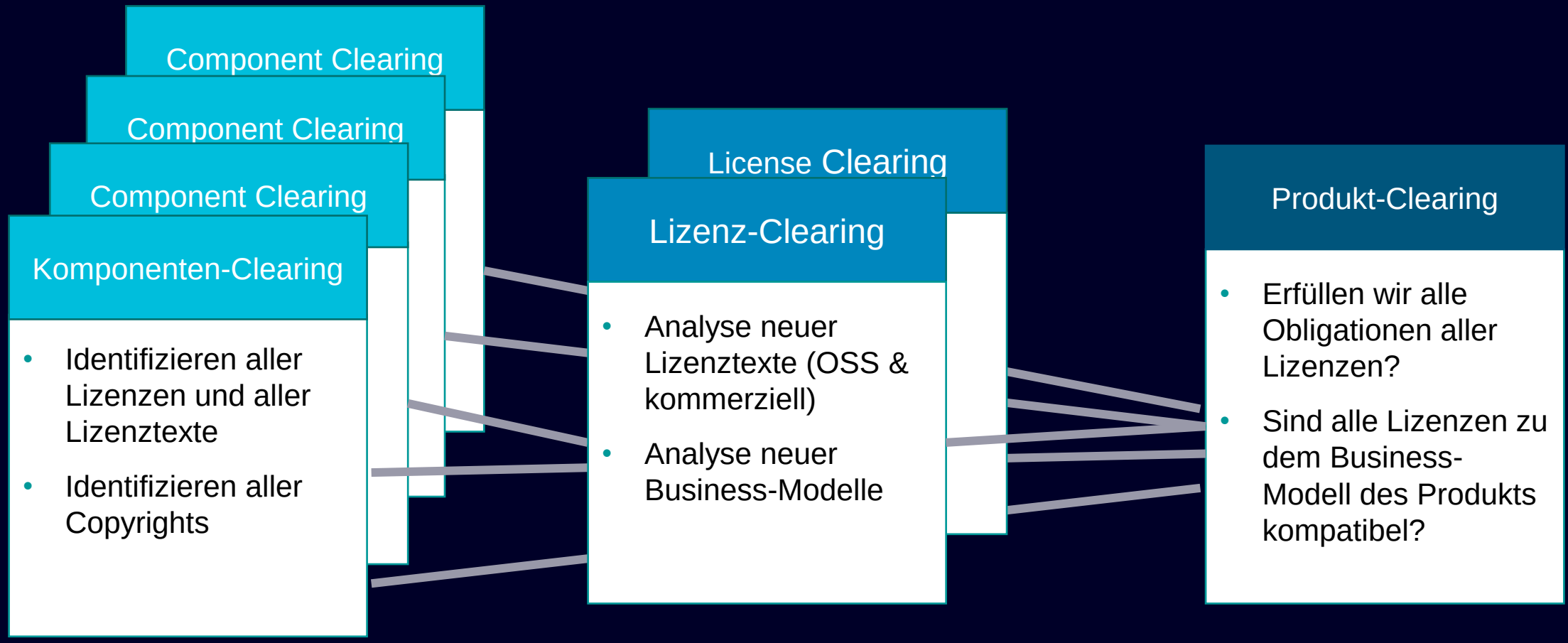
SIEMENS

SBOM und automatisierte Software License Compliance



SIEMENS

Verschiedene Arten von Clearing



SIEMENS

Automatische Identifikation von Softwarekomponenten

Funktioniert gut für

- Java, JavaScript, C#, Python, Go …

Funktioniert nicht gut für

- C/C++, embedded Software und Container

Und dann ist noch die Frage was ich zurückbekomme

- Alle Softwarekomponenten, d.h. auch die Softwarekomponenten die nur für das Bauen der Software benötigt werden (Build-Tools, Testframeworks)?
- Die Softwarekomponenten die tatsächlich benötigt werden?

Bootstrap hat keine echten Dependencies, aber mehr als 800 Development-Dependencies!

SIEMENS

Automatisches Finden von Metadaten

Wenn man Komponenten identifiziert hat, braucht man aber noch mehr Informationen: das Open Source Projekt, den Quellcode, die kommerzielle Lizenz…

Funktioniert gut für

- Java, JavaScript, C#, Python, Go …

Funktioniert nicht gut für

- Kommerzielle Komponenten, C, embedded Software und Container

Aber eigentlich ist das ja eine Aufgabe der Entwicklungsteams. Es ist ihre Software, also sollten sie doch genau wissen, was sie wo und aus welcher Quelle verwenden…

SIEMENS

Automatisieren des Eintragens in einen Software-Katalog

Würde gut funktionieren,

- wenn Namen eindeutig wäre
- wenn es eindeutige Identifier gäbe (package-url?)
- wenn GitHub-Projekte = Quellcode = Binaries (dll, jar, etc.)

Funktioniert gut wenn Menschen und Tools mitdenken!

Beispiel:

Apache Commons-Compress?
Apache Commons Compress?
Commons-Compress?
Commons Compress?
org.apache.commons.commons-compress?

SIEMENS

Automatisiertes Scannen nach Lizenzen und Copyrights

Die meisten Open Source Lizenzen fordern das Liefern des Lizenztextes und der Urheber

Dafür gibt es mehrere Möglichkeiten:

- Das Software-Ökosystem nach der (Haupt-)Lizenz fragen
- Einem Dienst vertrauen, der diese Informationen bereitstellt (kostenlos oder kommerziell)
- Jede Quellcode-Datei selbst scannen, z.B. mit FOSSology oder ScanCode

Einen sehr großen Teil der Arbeit kann man automatisieren, aber für exakte Ergebnisse braucht man dann doch immer wieder den Experten…

SIEMENS

Automatisiertes Scannen nach Lizenzen und Copyrights (2)

Beispiel:

```
// Use of this source code is governed by a BSD-style
// license that can be found in the LICENSE file.

// See:
//  - https://golang.org/LICENSE
//  - https://golang.org/src/crypto/x509/verify.go
```

Was macht man

- …wenn die Datei LICENSE nicht Teil des Quellcodes ist?
- …die URL nicht mehr funktioniert?
- …die Lizenzangaben unterschiedliche sind?

Automatisiertes Scannen nach Lizenzen und Copyrights (3)

Beispiel:

```
The C library (libtasn1.*) is licensed under the GNU Lesser General
Public License version 2.1 or later.  See the file COPYING.LIB.

The command line tool, self tests, examples, and other auxilliary
files, are licensed under the GNU General Public License version 3.0
or later.  See the file COPYING...

Part of the library are only these files:
* /lib/gl/limits.h …
```

… man muss also den Text und den Anwendungsfall verstehen …

SIEMENS

Automatisiertes Scannen nach Lizenzen und Copyrights (4)

Beispiel: jdk-jdk-19-36\src\java.smartcardio\unix\native\libj2pcsc\MUSCLE\COPYING

```
Some files are under GNU GPL v3 or any later version
  - doc/example/pcsc_demo.c
  - the files in src/spy/
  - the files in UnitaryTests/
```

GPL-3.0 in einem Produkt das wir auch an private Kunden liefern wäre ungünstig

…aber zum Glück sind alle genannten Dateien nicht Teil von OpenJDK 19-36.

Ein weit bekanntes kommerzielles Werkzeug kann das aber nicht erkennen…

Automatisierte Analyse von Lizenzen

Wenn man ein neue Lizenz findet, sollte man wissen

- was man darf (permissions), was man nicht darf (restrictions),
- was man tun muss (obligations),
- Was die Risiken sind (risks) und ob es Ausnahmen gibt (exceptions)

Dazu braucht man Anwälte … und man kann es bisher nicht automatisieren!

SIEMENS

Automatisiertes Aggregieren der Ergebnisse

Wenn man alle Komponenten und Lizenzen eines Produkts kennt, sollte man noch abschließend wissen, ob alles zum Use-Case des Produkts passt.

Bei Siemens nennen wir das Product Clearing.

Das ist jetzt ausnahmsweise etwas, was sich sehr gut automatisieren lässt:

- Welche Lizenzen gelten
- Muss ich auf einen möglichen Copyleft-Effekt achten
- Muss ich Vorbereitungen treffen, den Quellcode liefern zu können

SIEMENS

License Compliance für Container?

Lizenzbedingungen müssen eingehalten werden, daher …

- müssen wir alle Komponenten in allen Schichten des Images kennen
- müssen wir den Quellcode aller Open Source Komponenten haben
- müssen wir alle Lizenzen kennen, die in diesem Fall gelten

Wie liefere ich den Quellcode einer GPL-lizensierten Komponente, wenn meine Entwickler nicht einmal genau wissen was alles zu dem Image gehört, das sie verwenden?

SIEMENS

Container - Was die Experten sagen

Siemens hat die License Compliance Problematik mit international angesehenen Experten diskutiert:

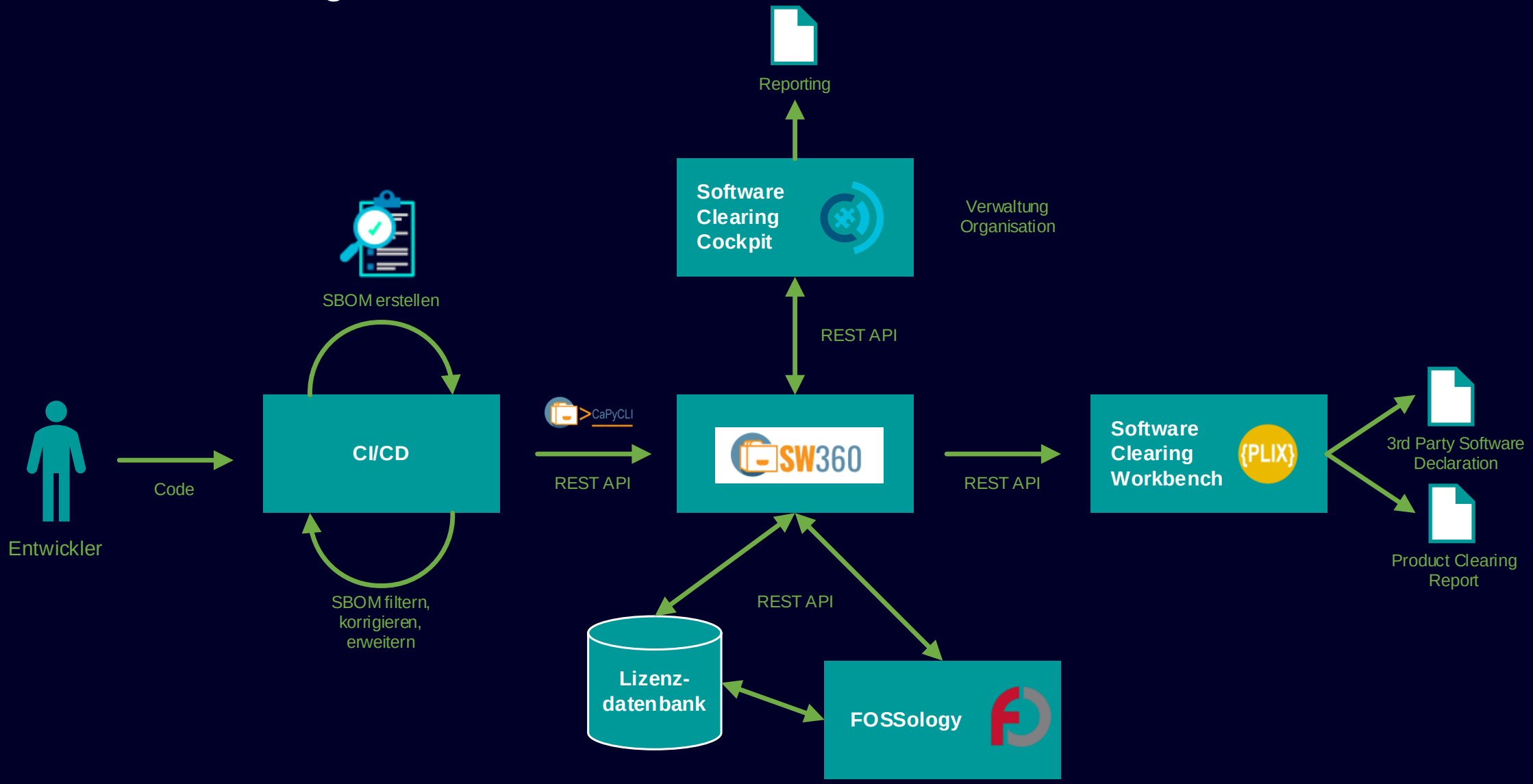
- Armijn Hemel (BAT), Philippe Ombredanne (ScanPipe), Rose Judge (Tern)

Aussagen:

- Container waren nur zum Testen gedacht und niemals zum Ausliefern von Produkten!
- Dockerfiles erlauben **alles …**
- Container können die Anzahl der verwendeten Third-Party Softwarekomponenten um den Faktor 10-100 erhöhen!
- **Die einzig zuverlässige Lösung: alle Container die man verwendet selbst bauen**

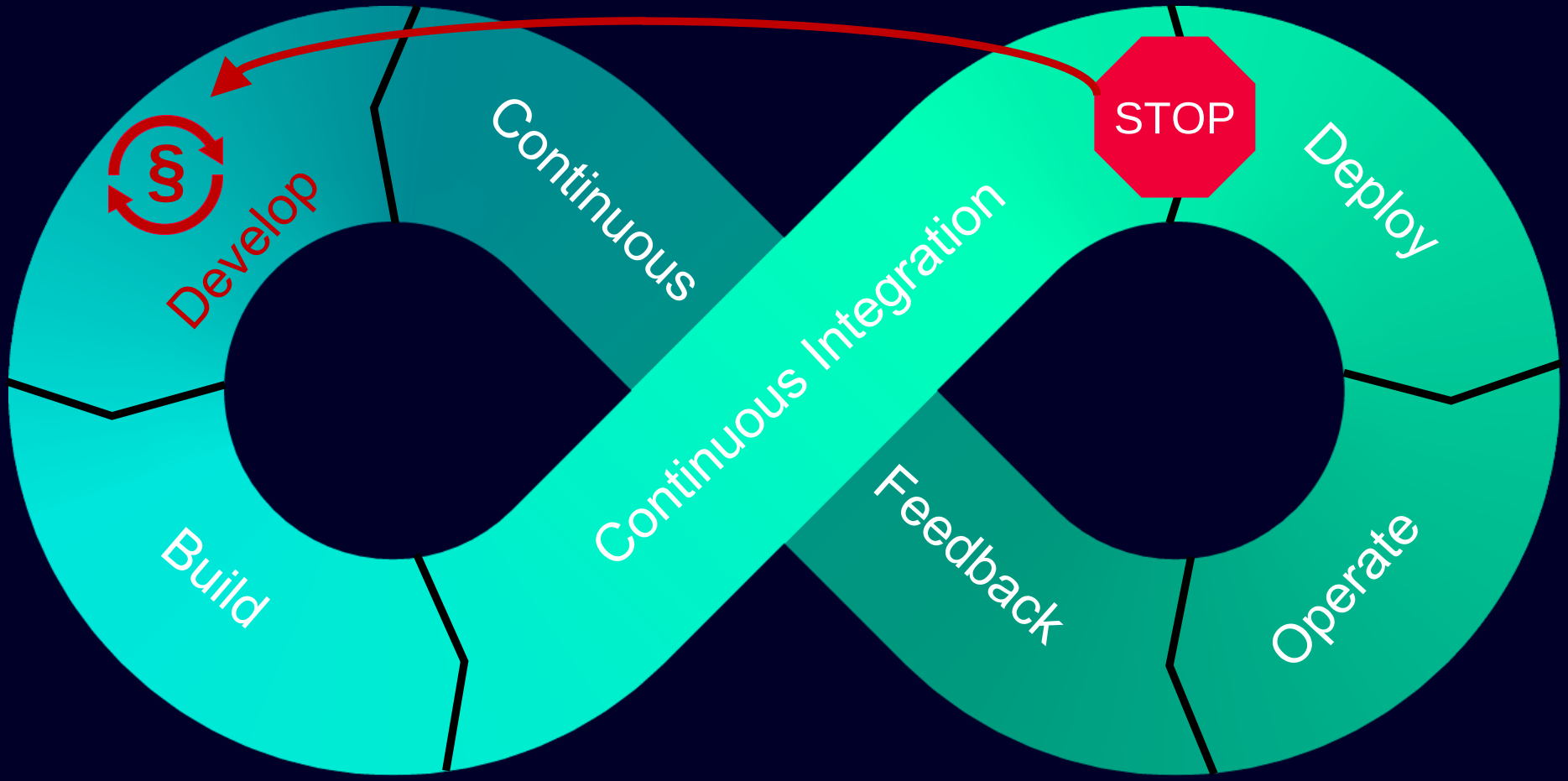
SIEMENS

Wie könnte das insgesamt aussehen



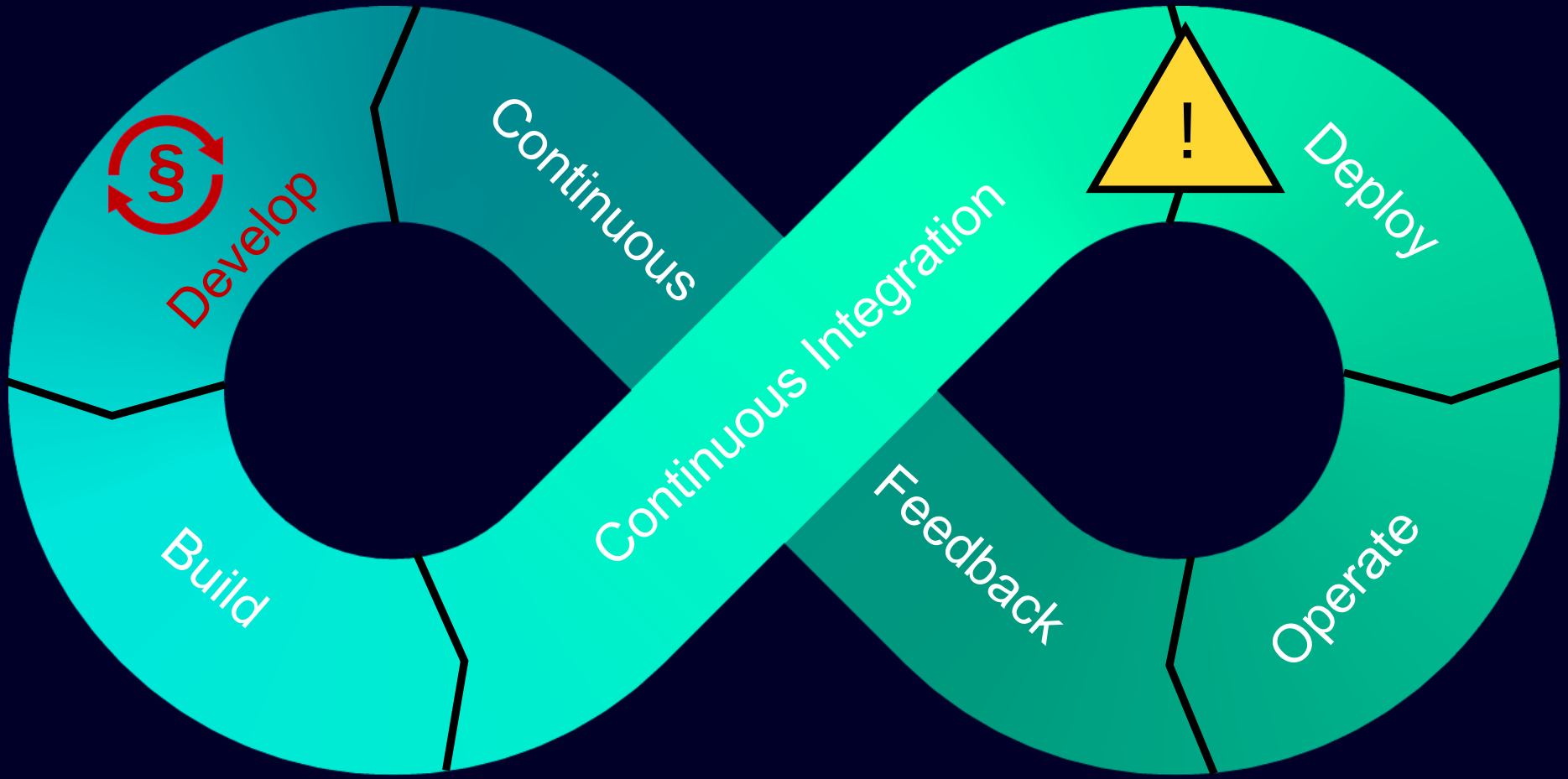
SIEMENS

CI/CD – Die Idee alles zu Automatisieren



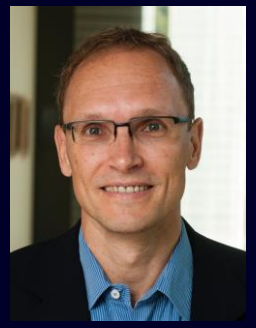
SIEMENS

CI/CD – Die Idee alles zu Automatisieren



SIEMENS

Contact



Thomas Graf
Principal Key Expert
Third Party Software License Manager
E-mail thomas.graf@siemens.com

Q&A

SIEMENS