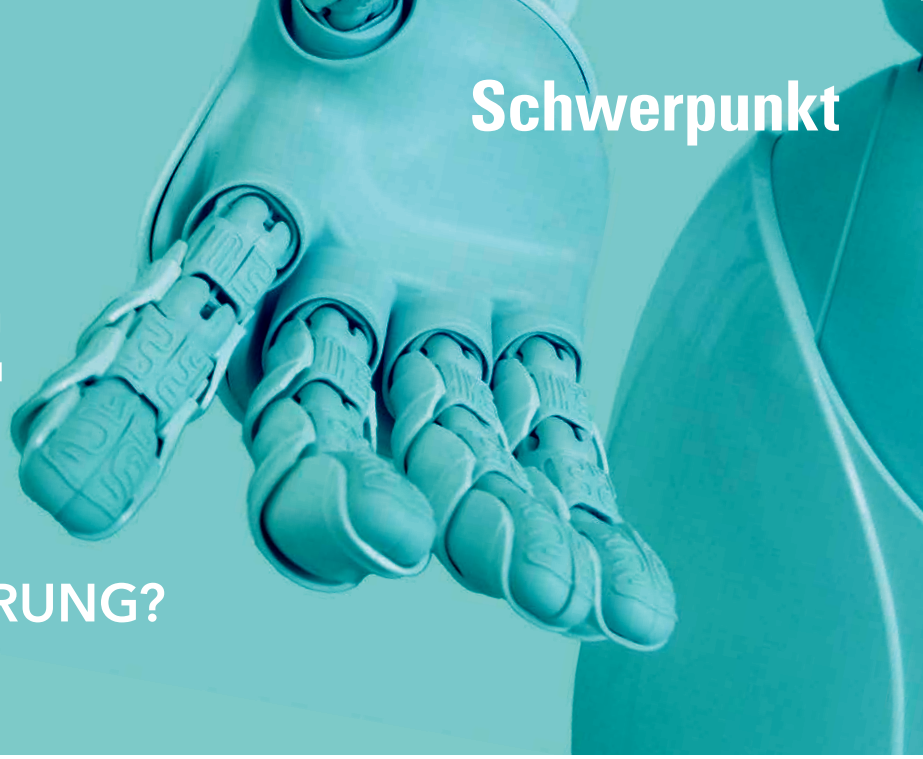

KI & SOFTWARE TESTEN

WIE TESTET MAN STÄNDIGE VERÄNDERUNG?

Methoden der Künstlichen Intelligenz sind älter als das Internet. Im Sommer 1956 fand die erste Konferenz zum Thema „Künstliche Intelligenz“ statt und bereits 1950 wurde der Turing Test entwickelt. Zur Künstlichen Intelligenz gehören neben den allgemein bekannten Neuronalen Netzen auch Fuzzy Logic, Maschinelles Lernen, Expertensysteme, Clustering und Klassifizierung sowie viele andere Methoden, die das Ziel haben, Entscheidungen auf Grundlage von komplexen Daten zu treffen. „KI könnte das schlimmste Ereignis der Menschheit werden“, so warnt Stephen Hawking vor den potentiellen Gefahren der KI am 9.11.2017 beim „Web summit“ in Lissabon, Portugal. Was gilt es bei der Entwicklung und vor allem dem Testen von KI zu beachten?

KI beruht zu einem großen Teil auf Software oder Software-Systemen. Diese KI-Systeme werden in verschiedene Hardware, wie Roboter, Kameras, Waschmaschinen oder Fahrzeuge, integriert. Wesentliche Qualitätsparameter wie die Sicherheit, Stabilität und Transparenz werden dabei zweitrangig behandelt. Aber besonders bei Systemen der Künstlichen Intelligenz (KI) sind die Folgen einer Manipulation nicht abzuschätzen. So gilt es nicht nur, die persönlichen Daten, sondern auch die KI vor Manipulation zu schützen.

In der Digitalen Transformation werden „Dinge“ intelligent. Die Vernetzung von realer und virtueller Welt lässt neue, disruptive Geschäftsmodelle entstehen und fördert diese zugleich. Damit trägt KI bereits heute maßgeblich zur Wertschöpfung in vielen Branchen bei. Dieser Trend wird weiter anhalten und entsprechende Risiken [1] mit sich bringen. So ist nicht nur mit weiteren Cyber-Attacken zu rechnen, sondern mit einem Black Out der IT-Systeme, der auf Grund der starken Abhängigkeit der Wirtschaft und Gesellschaft von funktionierenden IT-Systemen [6] einen Black Out des realen Lebens nach sich zieht. Diese Entwicklung enthält nicht nur technische, sondern auch ethische Risiken. Wie abhängig werden Mensch und Wirtschaft von der Künstlichen Intelligenz? Wie sehr werden Sie manipuliert und gesteuert? Wieviel Datenhoheit hat der Datenverursacher bzw. eigentliche Datenbesitzer überhaupt noch?
Bei der Untersuchung der verschiedenen Einsatzmöglichkeiten von Künstlicher Intelligenz in Systemen der Gesellschaft (z.B. Fuzzy Logik und Schwarmintelligenz bei der Balance zwischen Verbrauch und erneuerbarer Energieerzeugung [13]) werden die Risiken deutlich, die durch eine Manipulation eines solchen Systems existieren. Mögliche Szenarien werden deutlich in den Romanen „Black out“ und „Zero“ von Marc Elsberg beschrieben.

TYPEN VON KÜNSTLICHER INTELLIGENZ

Bei Künstlicher Intelligenz wird zwischen starker und schwacher Intelligenz unterschieden. Schwache künstliche Intelligenz fokussiert sich auf eine konkrete Anwendung, zum Beispiel um auf Grundlage von komplexen Daten Entscheidungen zu treffen. Beispiele hierfür sind Entscheidungen zur Kreditvergabe, Stellenbesetzung sowie Bild- und Spracherkennung, Übersetzungen, individualisierte Werbung und Navigationssysteme. Im Gegensatz dazu zeichnet sich starke KI durch logisches Denkvermögen, Entscheidungsfähigkeit auch bei Unsicherheit, Planungs- und Lernfähigkeit, Fähigkeit zur Kommunikation in natürlicher Sprache sowie Kombinieren aller Fähigkeiten zur Erreichung eines übergeordneten Ziels aus. Einige dieser Fähigkeiten – wie die Entscheidungsfähigkeit durch Unsicherheit – können bereits existierende KI-Systeme realisieren. Auch die Kommunikation in natürlicher Sprache ist schon weit entwickelt. Bei anderen Eigenschaften der Künstlichen Intelligenz fehlen zur Zeit noch allgemein anerkannte Definitionen und Regeln zum Nachweis der Fähigkeit.

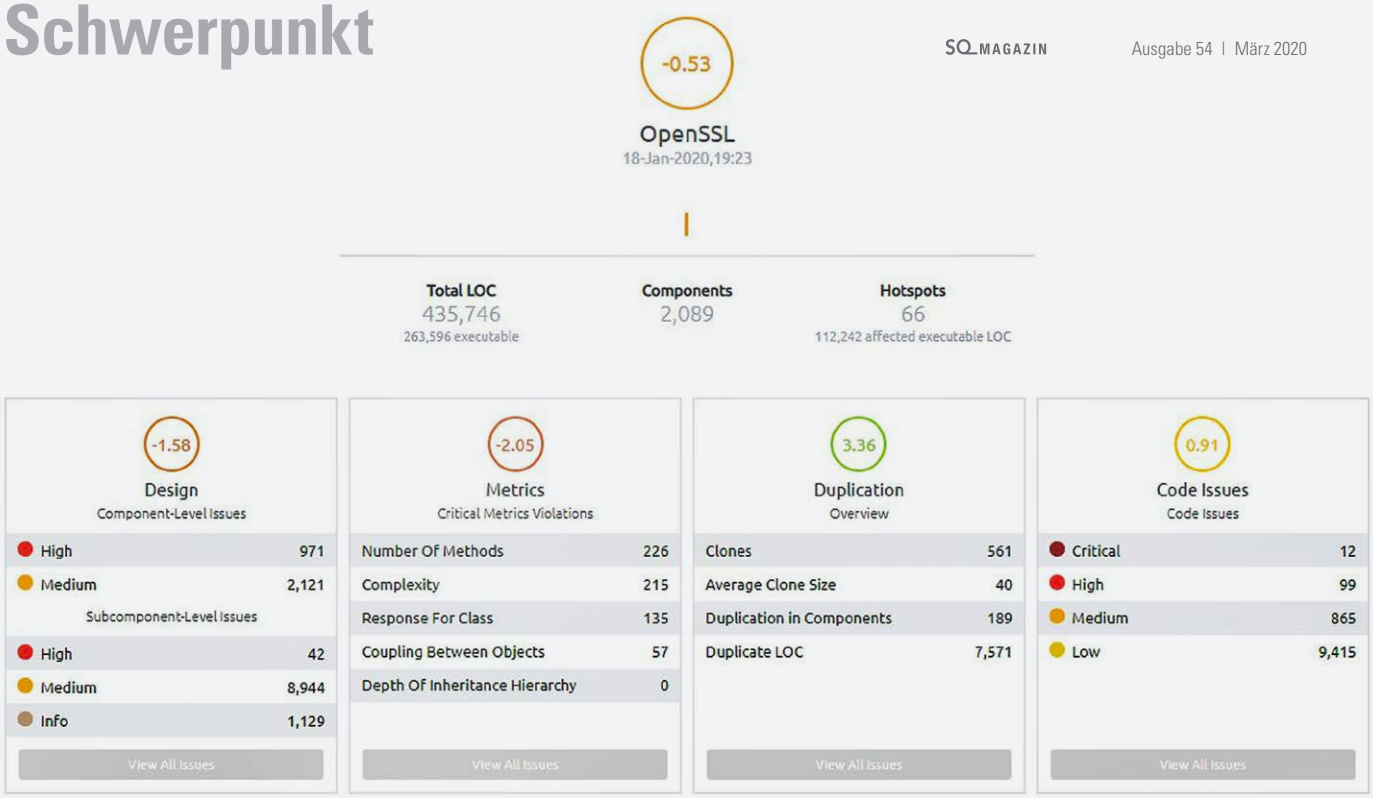

Abbildung 1: Code Analyse OpenSSL mit embold, 18.1.2020

Auf dieser Grundlage lassen sich vier Typen der Künstlichen Intelligenz definieren:

1 Reaktive KI, die durch Sensoren die Umwelt erkennt und entsprechend bestimmter Vorgaben reagiert und dabei keine Erinnerung bildet. Diese Systeme sind nicht in der Lage, Trends zu erkennen.

2 Systeme mit begrenztem Gedächtnis sind in der Lage, Erfahrungswerte zu speichern und auszuwerten. Sie können Trends erkennen und in die Entscheidungsfindung integrieren. Suchmaschinen, individualisierte Werbung, Chatbots und selbstfahrende Fahrzeuge sind Beispiele für diesen KI-Typ.

3 Systeme mit eigenem Bewusstsein erkennen sich und andere und verstehen Gedanken, Emotionen, Absichten, Motive und Erwartungen. Sie können sozial interagieren. Sonny aus dem Film „I Robot" ist ein Beispiel für eine derartige Künstliche Intelligenz.

4 Sich „ihrer selbst" bewusste Systeme können Vorstellungen über sich selbst bilden, sind sich ihrer inneren Zustände bewusst und können daraus Schlussfolgerungen ziehen und komplexe Strategien entwerfen. Sie können Gefühle anderer vorhersagen und in die eigenen Handlungen einbeziehen. Die Eva aus dem Film „Ex Machina" ist ein Beispiel für ein derartiges System.

HERAUSFORDERUNGEN FÜR DAS TESTEN VON KÜNSTLICHER INTELLIGENZ

Was unterscheidet KI von Software? KI erhält über Sensoren Informationen zur Umgebung. Sie kann – je nach Level oder Typ – auf Änderungen in der Umgebung reagieren und sich dabei an die Veränderungen anpassen. Grundsätzlich kann KI als komplexe „embedded Software" verstanden werden. **Daraus ergeben sich nun vier Herausforderungen für das Testen von KI:**

1. Komplexität

Die Komplexität von Software kann mit verschiedenen Metriken gemessen werden. Eine der bekanntesten Metriken ist die zyklomatische Komplexität (CC – Cyclomatic Complexity), oder auch WMC (weighted methods per class). Sie berechnen die Komplexitäten der Methoden einer Klasse. Viele Code Analyse Tools ermitteln die WMC oder zyklomatische Komplexität. So können die Klassen mit hoher Komplexität erkannt werden. Zur Orientierung wird zusätzlich ein Schwellwert angegeben, der nicht überschritten werden soll. Bei WMC liegt er je nach Berechnungsformel bei 17, bei CC (McCabe's method) bei 50 (toolabhängig und konfigurierbar). Je nach Anwendungsgebiet, Branche und Entwicklungsstand variiert die Komplexität einer Software. OpenSSL – eine bekannte Open Source Software mit 435.000 LoC (Lines of Code) – hat eine durchschnittliche Komplexität von 215, siehe Abb. 1, gemessen am 18.01.2020 mit embold, wobei die höchste Komplexität für eine Klasse 4.711 beträgt. OpenSSL ist ein Beispiel für eine umfangreiche und komplexe Software. Nicht jede KI-Methode erreicht diesen Umfang und diese Komplexität.

2. Entwicklungsgeschwindigkeit

Um Qualität und die Einhaltung von Standards zu sichern, muss KI wie jede andere Software vor der Inbetriebnahme getestet werden. Wenn KI sich selbst weiterentwickelt und Veränderungen in den Algorithmen vornimmt, werden die Release-Zeiten nicht mehr von einem Projektplan oder einem Standard wie Scrum bestimmt. Auch hier muss die Qualität vor der Inbetriebnahme getestet werden. Sobald die Software eine bessere Lösung für eine Aufgabe gefunden hat, werden die Algorithmen geändert. Bevor diese neuen Algorithmen aktiv werden können, müssen also