ATHENE
Nationales Forschungszentrum für angewandte Cybersicherheit

Offensive Cybersicherheit mit KI ist kein Orchideenthema

CPDP-Panel zum Thema „Offensive Cybersecurity by AI: Promises and Pitfalls“

AUTORINNEN

Indra Spiecker gen. Döhmann

Julia Esser

Caroline Plötz

Impressum
Nationales Forschungszentrum für angewandte
Cybersicherheit ATHENE
c/o Fraunhofer-Institut für
Sichere Informationstechnologie SIT
Rheinstraße 75
64295, Darmstadt

Hinweise
Dieser Beitrag wurde mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) und vom Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen ihrer gemeinsamen Förderung für das Nationale Forschungszentrum für angewandte Cybersicherheit ATHENE unterstützt.

Im Mai 2024 haben wir ein Panel zum Thema „Offensive Cybersecurity by AI: Promises and Pitfalls“ bei der renommierten Konferenz Computers, Privacy and Data Protection (CPDP) –dem europaweit größten Forum für aktuelle Fragen an der Schnittstelle von IT, Recht und Gesellschaft – veranstaltet. Die erfolgreiche Einwerbung des Panels sowie der Zuspruch des anwesenden Publikums zeugten davon, dass die Stärkung der offensiven Cybersicherheit ein Thema von zentraler, internationaler und interdisziplinärer Bedeutung ist. Die Diskussionsrunde wurde im Rahmen des ATHENE-Forschungsprojekts „Legal framework for offensive cybersecurity research (LegRes)“ organisiert von Caroline Plötz (Goethe-Universität Frankfurt a.M.) und Julia Esser (Universität zu Köln/Goethe-Universität Frankfurt a.M.), bzw. den Lehrstühlen Prof. Dr. Christoph Burchard (Goethe-Universität Frankfurt a.M.) und Prof. Dr. Indra Spiecker genannt Döhmann (Universität zu Köln/Goethe-Universität Frankfurt a.M.). Das Panel brachte hochkarätige Expert:innen aus Praxis, Forschung und Journalismus zusammen: Chris Kubecka (HypaSec), George Patsis (Obrela Security Industries), Dr. Johan Laux (Oxford Internet Institute) und Eva Wolfangel (freie Journalistin). Moderiert wurde die Diskussion von Prof. Dr. Indra Spiecker gen. Döhmann (Universität zu Köln/Goethe-Universität Frankfurt a.M.) und Dr. Annika Selzer (Fraunhofer SIT).

Die interdisziplinäre Expert:innenrunde bereicherte die CPDP 2024, die erstmalig auch KI-Themen integrierte, um eine hochaktuelle Debatte im Bereich der Cybersicherheit. Sie konnte die Öffentlichkeit mit einem Thema erreichen, welches trotz (oder vielleicht auch wegen) seiner Brisanz bislang wenig Aufmerksamkeit erhält. Aufgrund des stetig wachsenden Cyberspace muss der Einsatz von KI zur Vornahme offensiver Cybersicherheitsmaßnahmen als essenzieller Bestandteil der Cybersicherheitsarchitektur jedoch zunehmend in den Fokus rücken, weshalb die wichtigsten Punkte der Diskussion hier nochmal aufgegriffen werden:

Die Expert:innen bekräftigten den Befund, dass KI sowohl die Cyberbedrohungslandschaft als auch die offensiven Sicherheitspraktiken grundlegend verändert – Angriffe werden skalierbarer und präziser, aber auch weniger durchschau- und unmittelbar detektierbar. Sie bedrohen den Cyberraum, können und sollten aber auch zum Schutz der Cybersicherheit im Rahmen offensiver Cybersicherheitsmaßnahmen eingesetzt werden. Solche offensiven Sicherheitsmaßnahmen stellen einen Analyseansatz zur Testung von IT-Systemen dar. Sie bedienen sich Methoden, die denen maliziöser Hacker:innen ähneln oder gar mit diesen übereinstimmen. Anhand der vorgenommenen Angriffssimulationen können proaktiv Schwachstellen in IT-Systemen aufgedeckt und in der Folge die Sicherheit der IT-Infrastruktur gestärkt werden. Die Chancen des KI-Einsatzes in der Cybersicherheit und den aktuellen technischen Stand verdeutlichte die ethische Hackerin Kubecka anhand ihres selbst entwickelten KI-Modells zur Schwachstellenanalyse, das ihr zahlreiche Arbeitsschritte abnehme. Einigkeit bestand darin, dass Cyberangriffe und deren Simulation unter Verwendung von KI längst keine Zukunftsmusik mehr, sondern im IT-Alltag angekommen sind.

Wie lässt sich aber Verantwortung für die mit den Angriffen einhergehenden Datenzugriffe und verwirklichten Straftatbestände zuweisen, wenn die Testsysteme zunehmend eigenständig agieren? Hinter dieser Frage steckt eine grundlegende Debatte zur Verantwortungszuweisung beim Einsatz von KI, die – so zeigte auch die Diskussion im Panel – hochaktuell und keinesfalls abgeschlossen ist. Laux verwies auf die besondere regulatorische Bedeutung menschlicher Kontrolle („human oversight“). Eine solche Kontrolle erfordert allerdings die Nachvollziehbarkeit der KI-Aktion, deren Umsetzbarkeit die Expert:innen im Panel unterschiedlich einschätzten: Wolfangel etwa bezweifelte diese, während Kubecka die Dokumentation KI-basierter Entscheidungswege durchaus für möglich hielt. Sie sei aber mit zusätzlichem Aufwand verbunden.

Spiecker gen. Döhmann führte an, dass daneben die Frage nach der Wahrhaftigkeit zu beantworten sei – wie könne sichergestellt werden, dass die von der KI selbst abgegebene Erklärung zu ihrer Entscheidung auch der Wahrheit entspricht? Aus Sicht des Praktikers Patsis sei das keine Frage der Technik, sondern eine solche des Vertrauens. Wie etwa bei Ärzt:innen müsste auch gegenüber KI-gestützten Systemen durch Auditierung, Zertifizierung und Erfahrungswerte Vertrauen hergestellt werden. Dass dies über die KI-VO aber nicht hinreichend geleistet werde, war in der Diskussion offenkundig.

Die Regulierung offensiver Cybersicherheit wurde kontrovers diskutiert. Einigkeit bestand zunächst darin, dass eine dezentrale Strukturierung von KI-Systemen einen Gewinn an Sicherheit bedeute, allerdings weise der aktuelle Rechtsrahmen Lücken auf, so Spiecker gen. Döhmann und Kubecka. Patsis hingegen hielt die bestehende Regulierung für grundsätzlich geeignet, die offensive Cybersicherheit abzubilden, verwies aber ebenso wie Laux auf die regulatorische „Technologieagnostik". Das Recht verschließe sich technischen Neuerungen und passe sich nicht entsprechend an. Deutlich wurde, dass sich Recht und Technik nicht im Gleichschritt entwickeln – die Frage, wie und ob dieses Problem zu lösen sei, blieb offen und wird zukünftig zu beleuchten sein.

Die Diskussion zeigte, dass eine werteorientierte und risikobewusste Technikgestaltung notwendig, jedoch keinesfalls hinreichend ist. Die Technik allein – hier konkret die Anbindung von KI in offensive Cybersicherheitsmaßnahmen – genügt nicht zur Sicherstellung einer funktionierenden Cybersicherheitsarchitektur. Bislang sehen sich KI-Nutzer:innen im Bereich der offensiven Cybersicherheit u.a. mit ethischen Fragen konfrontiert, für die es keine Standards – und oftmals keine Handlungsanweisungen – gibt. Zudem laufen sie beim Einsatz von KI-gestützten offensiven Methoden verstärkt Gefahr, gegen rechtliche Vorgaben zu verstoßen und sich so haftbar oder strafbar zu machen.

Der Einsatz von KI zu Zwecken der Angriffssimulation ist längst kein Orchideenthema mehr, sondern ein drängendes Handlungsfeld – wie auch die abschließende, sehr lebhafte Diskussion mit dem anwesenden Publikum zeigte. Zukünftig bedarf es deshalb eines interdisziplinären Austauschs zwischen Forschung und Praxis, um die Lücke zwischen Technik und Recht zu schließen, sowie differenzierte Regulierungen und Handlungsempfehlungen zu entwickeln.