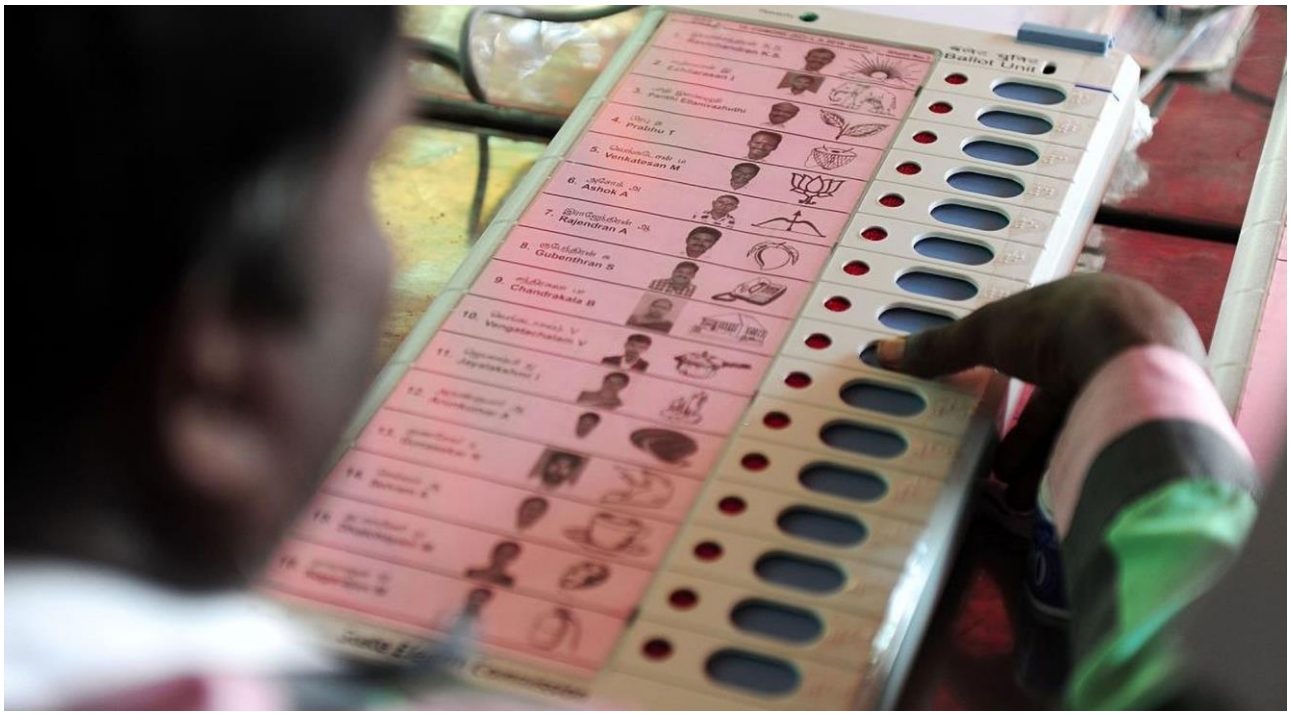
Januar 2021

Länderbericht

New York

KONRAD ADENAUER STIFTUNG



Cyber-KI Konvergenzen und ihre Wirkungen

Die Sicherheit von Wahlprozessen und deren Schwachstellen

Kurzfassung der Publikation "Cyber-AI Convergence and Interference: Securing elections and building human resilience" von Eleonore Pauwels

Wir stehen vor neuen Herausforderungen in Bezug auf die Sicherheit von Bürgern und in der Politik, die weit darüber hinaus gehen, was traditionelle Sicherheits- und Militärdoktrinen bislang umfasst haben. Wichtig ist heute nicht nur die Frage, wer neue Gebiete erobert, sondern wer die Daten, das Vertrauen, die „Hearts and Minds" der Bürger eines Landes oder eines politischen Systems für sich gewinnt.

Seit einem Jahrzehnt nutzen fremde Mächte mit unlauteren Absichten die Infrastruktur, die demokratische Gesellschaften untermauert, als Tatwerkzeug. Hacker sind in das Internet, in Medien und sogar in Wählerdatenbanken eingedrungen, um Verwirrung zu stiften, Unzufriedenheit und Misstrauen zu säen. Beim Brexit-Referendum 2016, in den Vorwahlen zur US-Präsidentschaft wie auch in den eigentlichen Wahlen 2016, aber auch bei der Wahl des französischen Präsidenten 2017 haben
fremde Störer systematisch versucht, den demokratischen Diskurs zu verzerren.

“Eine neu entstehende Typologie von Cyberangriffen könnte sich KI zunutze machen, um die Integrität der am Wahlprozess beteiligten Datensätze und Softwareprogramme zu manipulieren“

Sowohl staatliche als auch nichtstaatliche Akteure nutzen bereits die Konvergenzen von künstlicher Intelligenz (KI) und Cyber-Fähigkeiten, um Informationen zu manipulieren, Vertrauen zu untergraben, und um in die internen politischen Prozesse anderer Staaten einzugreifen oder Infrastrukturen lahmzulegen, die für die nationale und persönliche Sicherheit entscheidend sind.
Diese technologische Konvergenz hat bedeutende nachteilige, soziale und sogar strategische Auswirkungen. So könnte die Nutzung von KI-Systemen beispielsweise den Charakter, die Tragweite und Intensität von Cyberangriffen auf die kritische Wahlinfrastruktur der Staaten erheblich verstärken.

Eine neu entstehende Typologie von Cyberangriffen könnte sich feindliche KI („adversarial AI“) zunutze machen, um die Integrität der am Wahlprozess beteiligten Datensätze und Softwareprogramme zu manipulieren. Derartige feindliche Angriffe setzen schon jetzt Methoden ein, um sich zu tarnen, menschliche Schwachstellen durch präzises Social Engineering gezielt aufzudecken und letztendlich die Cyber- und Informationssicherheit zu schädigen. David Schwed, Professor, Gründer und Leiter des Programms für Cybersecurity an der Katz School der Yeshiva University betont: „KI wird eine zunehmend wichtigere Rolle im Angriffsarsenal der Gegner spielen. Sie werden in der Lage sein, unbegrenzte autonome Angriffe zu lancieren, mit geringerem Bedarf an menschlichen nachrichtendienstlichen Informationen.“[1]

Staaten werden lernen, mit diesen Cyber-Bedrohungen für ihre Wahlen umzugehen, genauso wie sie die sich verändernde Art und Tragweite von Cyberkonflikten mit niedriger Intensität verstehen lernen. Die Hauptsorge wird dabei bleiben, dass diese Bedrohungen für die Wahlsicherheit durch KI und die zunehmende Cyber-Vernetzung komplexer werden und sich schwerer erkennen und verhindern lassen.

Angriffe werden auf die nationale Informationsinfrastruktur abzielen und damit die Integrität sensibler Sicherheitsdaten und biometrischer Informationen der Bürger untergraben. Mit einer zunehmenden Anzahl der an das Internet angeschlossenen Geräte – von persönlichen Sensoren bis zu Elementen wichtiger Infrastruktur – wird der Einsatz von 5G die KI-Bearbeitung am Endgerät beschleunigen. Dadurch erhöhen sich die Gelegenheiten für und die Zerstörungskraft von raffinierten Cyberangriffen auf Wahlen.

„Diese Bedrohungen für die Wahlsicherheit werden durch KI und die zunehmende Cyber-Vernetzung komplexer, schwerer zu verhindern und zu erkennen“

Die vorliegende technische und politische Analyse untersucht, wie die Herausforderungen bei der Nutzung von KI und Cybertechnologien, in den umfassenden Strategien zur Unterstützung von Wahlprozessen adressiert werden,[2] einschließlich der Strategien von Organisationen wie IFES[3] oder EU-Mitgliedsstaaten.[4] Die vorliegende Untersuchung befasst sich besonders mit der zunehmenden Anfälligkeit digitaler Tools zur Verwaltung und Unterstützung von Wahlprozessen gegenüber KI-gestützter Malware und Cyberangriffen und verweist auf den Bedarf nach neuen Ansätzen zur Sicherung und Stärkung der Widerstandsfähigkeit der Wahlinfrastruktur. Die Fähigkeit autonomer Malware, ihre eigenen Strategien zu verbessern und mit jedem Durchlauf immer aggressivere, präzisere Gegenangriffe zu lancieren, führt zu einer Erweiterung und Steigerung vorhandener Fähigkeiten für Cyberangriffe. Die Automatisierung von Cyberangriffen, mit denen die Integrität kritischer Informationen innerhalb der Wahlinfrastruktur manipuliert und beschädigt werden kann, ist eine wachsende Bedrohung, die durch die technologische Konvergenz ausgelöst wird. Dieser Bericht zeichnet ein Bild der neu entstehenden Cyber-Bedrohungen und analysiert den Wahlzyklus und die Infrastruktur, um Einstiegspunkte für konvergierende Cyber-KI-Angriffsvektoren zu ermitteln, dazugehörige Datenziele und Schwachstellen zu erkennen und Empfehlungen auszusprechen.

Auf der Basis von anderen grundlegenden Untersuchungen[5] wählt die Autorin einen spezifischen Ansatz, der den Wahlzyklus und seine Infrastruktur zunächst als eine Gruppierung komplexer sozio-technischer Systeme und in zweiter Linie als einen Satz von datengestützten Prozessen betrachtet. Dieser Ansatz ist ganzheitlich und strategisch, da er erlaubt, neue Schwachstellen, die KI zunehmend innerhalb der Dateninfrastrukturen und Datenoptimierungsprozesse bei der Durchführung von Wahlen angreifen kann, vorauszusagen, einzuordnen und damit besser zu verstehen. Der Autorin geht es weniger darum, das Ausmaß der Digitalisierung einzelner Phasen im Wahlprozess zu thematisieren, sondern darum, Bedrohungen für die Datenintegrität im gesamten Informationsverarbeitungsszyklus eines Wahlprozesses zu prognostizieren.

Manipulation der Datenintegrität ist eine neue und äußerst wirksame Taktik all jener, die für relevante sozio-technische Systeme Betrugsverdacht und Misstrauen schüren wollen. Wahlen – wie auch andere wichtige datengestützte Infrastruktur im Gesundheitswesen oder in der Katastrophenhilfe – sind anfällig für neu auftretende Methoden der Datenmanipulation und -fälschung.[6] Wie das Vertrauen in Gesundheitsdienste und Katastrophenmanagement, so stellt auch das Vertrauen in Wahlen einen Kernbestandteil unseres Sozialvertrags dar. Mehr noch, es ist die Grundlage für unsere Demokratien.

„Die Automatisierung von Cyberangriffen, mit denen die Integrität kritischer Informationen innerhalb der Wahlinfrastruktur manipuliert und beschädigt werden kann, ist eine wachsende Bedrohung, die durch die technologische Konvergenz ausgelöst wird."

Im Kontext der Sicherheit von Wahlen, wo es entscheidend ist, Vertrauen aufzubauen und zu stärken, bringt die Präsenz von KI sowohl eine erkenntnistheoretische als auch eine neue Technologieebene. KI-Technologien geben vor, uns dabei zu helfen, einen komplexen Wissensfundus zu produzieren, zu analysieren, geltend zu machen und zu verifizieren. Gleichzeitig ermöglicht es diese Technik aber auch die Integrität und Glaubwürdigkeit unserer globalen Informations- und Informationsgewinnungssysteme zu untergraben.[7]

Abschnitt 1 des vorliegenden Berichts beschreibt die jüngsten Trends und Forschungen in den Bereichen Cybersicherheit und Wahltechnologie und bietet eine erste Diagnose, wie sich traditionelle Legacy[8]-Ansätze in der wahlbezogenen Cybersecurity an das Zusammenfließen von Cyber- und KI-Technik anpassen müssen.

Abschnitt 2 formuliert ein Paradigma der KI-Konvergenz und erklärt, wie dieses Paradigma dazu beiträgt, KI gestuetzte Cyberangriffe zu ermoeglichen, die wiederum in der Lage sind, die Integrität von Datensätzen gezielt anzugreifen und zu manipulieren.

Abschnitt 3 stellt eine allgemeine Matrix eines Wahlzyklus sowie einer Wahlinfrastruktur bereit und analysiert das Spektrum konvergierender Bedrohungen und Schwachstellen, Angriffsvektoren, Datenziele und Auswirkungen. Abschnitt 4 schließlich enthält Empfehlungen, insbesondere Überlegungen zu den Anforderungen und Methoden für eine Bedrohungsanalyseund zur Rolle und Verantwortung unterschiedlicher Stakeholder.

1 Forbes, 2019. „141 Cybersecurity Predictions for 2020." https://www.forbes.com/sites/gilpress/2019/12/03/141-cybersecurity- predictions-for-2020/

2 Dieser Bericht beruht auf primären sowie sekundären Ressourcen und verfolgt einen Ansatz aus kombinierten Methoden, darunter qualitative Schreibtischstudien, Fachliteraturrecherchen, Politikanalysen, Befragungen von und Beratungen mit Fachleuten sowie vorausschauende Methoden (Signale, Treiber, Trend-Impact-Analysen). Die Liste der primären und sekundären Ressourcen findet sich im Anhang.

3 International Foundation for Election Security (IFES), 2018. „Cybersecurity in Elections." https://www.ifes.org/sites/default/files/2018_heat_cybersecurity_in_elections.pdf

4 NIS Cooperation Group, 2018. „Compendium on Cyber Security of Election Technology." https://www.ria.ee/sites/default/files/content-editors/kuberturve/cyber_security_of_election_technology.pdf

5 NIS Cooperation Group, 2018.; IFES, 2018.; Herpig S., et al, 2018. „Securing Democracy in Cyberspace." https://www.stiftung- nv.de/sites/default/files/securing_democracy_in_cyberspace.pdf

6 Datenvergiftung ist ein feindlicher Angriff, der darauf abzielt, den Trainingsdatensatz zu manipulieren, um das Vorhersageverhalten eines trainierten Algorithmusmodells so zu kontrollieren, dass das Modell feindliche Beispiele („adversarial examples") in eine erwünschte Kategorie einordnet (z. B. Spam-E-Mails als sicher einstuft). Datenvergiftungs-Angriffe können somit den Lernprozess eines Machine Learning-Systems unterwandern und/oder die Leistung des Systems herabsetzen.

7 Pauwels E., 2019. „The New Geopolitics of Converging Risks." https://collections.unu.edu/eserv/UNU:7308/PauwelsAIGeopolitics.pdf

8 Im Rahmen dieser technischen und politischen Abhandlung über KI und Cybersicherheit beschreibt der Begriff „Legacy" ein System oder einen Ansatz, der alt, traditionell („vererbt") ist, aber noch als Referenz benutzt wird, weil er zu schwierig zu ersetzen wäre. Die zum Aktualisieren und Ersetzen eines solchen „Legacy-Ansatzes" erforderliche Kenntnis und Weitsicht wurde noch nicht erworben und erreicht.

Konrad-Adenauer-Stiftung e. V.

Andrea Ostheimer
Executive Director
New York
www.kas.de/newyork

andrea.ostheimer@kas.de

Photo: © Arun Sankar / AFP

The text of this publication is published under a Creative Commons license: "Creative Commons Attribution- Share Alike 4.0 international" (CC BY-SA 4.0), https://creativecommons.org/licenses/by-sa/4.0/legalcode